Drug side-effect prediction using machine learning methods

Muhammad Irfan Khan

School of Science

Thesis submitted for examination for the degree of Master of Science in Technology.

Espoo November 24, 2017

Supervisor

Prof. Samuel Kaski , Dr. Murugan Natarajan Arul

Advisors

Dr Pekka Martinen

Dr Jing Tang

Aalto University
School of Science

Aalto University
School of Science

Aalto University, P.O. BOX 11000, 00076 AALTO
www.aalto.fi
Abstract of the master's thesis

Author Muhammad Irfan Khan

Title Drug side-effect prediction using machine learning methods

Degree programme euSysBio

Major Computational Systems Biology **Code of major** T-61

Supervisor Prof. Samuel Kaski , Dr. Murugan Natarajan Arul

Advisors Dr Pekka Martinen, Dr Jing Tang

Date November 24, 2017 **Number of pages** 52+1 **Language** English

Abstract

Drug toxicity (or adverse side effects) is a pressing health problem which is also an impediment to the development of therapeutically effective drugs. Despite many on-going efforts to determine the toxicity beforehand, computational prediction of drug side-effects remains a challenging task.

This thesis presents an approach to predict side-effects by utilizing side-information sources for the drugs, while simultaneously comparing state-of-the-art machine learning methods to improve accuracy. Specifically, the thesis implements a data-analysis pipeline for obtaining side-information that are useful for the prediction task. This thesis then formulates the drug side-effect prediction as a machine learning problem: Given disease indications and structural features (as side-information sources) of drugs, for which some measurements of side-effect exist, predict side-effect for a new drug.

As case studies, the prediction accuracies are compared for ten different side-effects using linear as well as non-linear machine learning methods. The thesis summarizes three key findings. First, the drug side-information sources are predictive of the side-effects. Second, non-linear methods show improved prediction accuracies as compared to their linear analogs. Third, the integration of disease indications and structural features with a principled machine learning approach further improves the drug side-effect predictions.

However, the current study limits the analysis assuming side-effects are independent. In future, modeling the joint relationships of several side-effects could yield more strong predictions and better help to understand the underlying biological mechanism.

Keywords Machine Learning, Side-effect Prediction

Preface

The reported work constitutes my Master Thesis and concludes my Master Degree within the Erasmus Mundus Master program in Systems Biology (euSysBio), a joint program KTH and Aalto University.

This work has been carried out in Quantitative Systems Pharmacology group in Institute for Molecular Medicine Finland, University of Helsinki in collaboration with Probabilistic Machine Learning(PML) group in Department of Computer Science, Aalto University School of Science, Finland. This thesis has been funded by HiLIFE Research Fellow program.

I am grateful to the Erasmus Mundus program for giving me the opportunity to pursue my studies. My sincere complements to my supervisors and instructors for training me and teaching me the principles of science and for the opportunity to work in interdisciplinary field.

I am also thankful to Dr. Ammad-ud-din, Dr. Suleiman Ali Khan, Dr. Zia-ur-Rehman, Alina Malyutina, Zaid Alam, Dr. Rao Anwar, Wahaj-ud-din and Mansoor khan for there guidance, fruitful discussions and feedback.

Finally, I am indebted to my parents for their support and blessings throughout my life. I thank my parents and siblings for instilling in me the value of education.

Otaniemi, 24.11.2017

Muhammad Irfan Khan

Contents

List of Figures

List of Tables

Symbols and abbreviations

Abbreviations

BEMKL	Bayesian efficient multi kernel learning
SVM	Support Vector Machines
LDA	Linear Discriminant Analysis
LR	Logistic Regression
NN	Neural Networks
RF	Random Forest
NB	Naive Bayes

1 Introduction

Drug side-effects/ adverse drug reactions (ADRs) is a crucial and complex challenge. The research community is concerned as drug toxicity is the fourth leading cause of death in U.S alone after cancer and heart diseases (Leone et al., 2008)(Bloomquist, n.d.). Moreover, If the drug success rate in clinical trials increases from 25 percent to 33 percent, pharmaceutical companies can save around 200 million dollars on the drug development process and reduce one fourth of the total drug development time (DiMasi, 2002). Effective ADRs prediction is essential for improving patients healthcare and accelerating the drug development process.

Different computational techniques have been used in recent past in order to understand the mechanism of drug reactions. The data sources used to study side-effect in different studies include chemical data of both drugs and drug-targets. A literature review of these cutting-edge approaches is summarized in section 1.2 of this thesis. The major cause of drug side-effect is off-target reactions. The mechanism of action of drugs is influenced by the genomic heterogeneity of individuals and influencing chemical properties due to altering micro-environment in cellular compartments. Hence, side-effects "as clinical phenotypes" that arise in patients can assumed to be a manifestation of complex interaction of multitude of factors i-e genomic features, disease state in which drugs are administered called drug indications, chemical descriptors of drugs (Schuster, Laggner, & Langer, 2008).

Unlike drug chemical and structural properties, the genomic information of the patients is not available from public databases. A disease or a group of specific biological symptoms arises due to abnormal "interaction" or "cross-talk" of pathways. Drugs are usually administered to restore the normal state by triggering the cascade of reactions in perturbed pathways. Here genome plays a fundamental role in mechanizing biological reactions of these drugs; moreover, similar drugs are given for similar diseases. Therefore, this thesis considers that the drug indications (also known as drug-disease associations) are an approximation of missing genomic information from the patients.

The basic research question this thesis aims to solve is if the chemical discriptors called as fingerprints of the drugs and drug indications are an insightful information source for the drug side-effects. Finding out these underlying relations of drugs side-effect that arise due to this complex interaction is the motivation behind this study. In this thesis, drug indications and chemical descriptors are used as data sources to predict ten different "common" side-effects. An analysis pipeline to predict the general side-effects is developed for this research work. A straightforward comparison of seven different machine learning algorithms and their prediction performance is presented as a result. The selected side-effects are grouped together in two categories. The categories are based on which data source is more insightful for that particular side-effect prediction.

In this thesis, computational analysis is performed to find a link between biological responses of patients and chemical properties of drugs. The study is constituted on three disciplines: pharmacology, chemi-informatics and machine learning. The aim is to systematically examine the relevant publicly available data for effective

modelling of drug side-effects. Comparison of prediction performance of different machine learning models is made, while simultaneously co-behaving side-effects (similar behaviour of machine learning approaches on data type) are grouped together in an attempt to better learn the pattern in interactions.

1.1 Background

After drug assembly, the life cycle of drug development is comprised of two major stages, pharmacokinetics and pharmacodynamics. Pharmacokinetics can be divided in many levels as absorption and assimilation (what is the intake route for drug for maximum affect), distribution (how is the drug molecule transfered in biological system i-e cell), metabolism (what is the mechanism of action adopted by the drug and how is it processed in the biological system) and elimination (how it is excreted out of the system). Drugs can be administered through different routes including ingestion, intravenous or intramuscular uptake. Drugs are chemical entities which interact with biomolecules to trigger a reaction which leads to an altered cellular state. Drugs can interact with targeted biomolecules of interest or unintended off-target biomolecules as well which can cause undesired affects on subjected patients. These undesired adverse affects are often classified as drug side-effects, toxicological properties of drugs or Adverse drug reactions (ADR) profiles of pharmaceutical agents.

Drug side effect studies are a part of pharmacodynamics i-e the impact of concentration of drug substance on various organs of the body over a monitored course of time. Studies of side effects of drugs are significant because drugs which are cytotoxic can be rather harmful than beneficial to the patients. (Kola & Landis, 2004). Ultimately, clinical safety is the standard and the most stringent checkpoint for all drugs to be approved by Food and Drug Administration (FDA). Practically, successful clinical trial is the most significant step in life cycle of drugs development which contributes to translational medicine. Safety and efficacy of drugs and vaccines is the main mission statement of pharmacology industry and benefactor to improved healthcare system in general.

Doctors and clinicians recommend the drugs which have minimum reported side-effects by patients. Pharma Industry (R&D) continuously investigates and work towards designing and bettering the therapies.(Terstappen & Reggiani, 2001). Demand for the new and improved, effective and efficient drugs are always on the high, but the side-effects caused often due to off-target effects of the drugs is a major problem in medicinal science. Traditionally, side-effects of drugs have been observed and recorded in pre-clinical trials on different animal models (i-e lab rats, chimpanzees or volunteers). Moreover, results from animal models are not generalizable across species (Bracken, 2009) .

Besides, there is high risk of harmful toxic effects which could be far-reaching (potentially fatal in some cases) and injurious to health of volunteers undergoing investigational therapy if testing drug is of addictive nature as well. This practice is hazardous and the overall process is time-consuming. If the drug success rate in clinical trials increases from 25 percent to 33 percent, pharmaceutical companies can

save around 200 million dollars on the drug development process and reduce one fourth of the total drug development time (DiMasi, 2002).

Higher efficacy of drugs lead to better treatment plans which can be ensured if the sideeffects of the drugs can be prevented. In order to ensure prevention of side-effects, prediction of the side-effects from available multiple different data sources is required. Recent technological advances has enabled the integration of clinical observation and molecular biology data. It is a relatively new approach and it is called Systems Pharmacology, and there are very limited studies for prediction of side-effects of drugs using experimental drugs data (Huang, Wu, & Chen, 2011). With increasing computational processing power and robust machine learning models, it is possible to predict the side-effects of the putative drugs and accelerate the development of drug-design process at large (Ma et al., 2008) (T.-B. Ho, Le, Thai, & Taewijit, 2016).

To understand the impact of drugs (chemical entities) on patients is an important task in medical science. Predicting adverse effects of drugs is valuable for improving drug design and targeted therapies as this is the most basic screening method for drug design.(Chiang & Butte, 2009) The economic damage caused by drug-side-effect problem is highlighted by the fact that approximately 30 percent drugs do not proceed beyond trial phase in clinics due to adverse side-effect and drug efficacy (J. Zhang, 2012). In United States, estimated 100,000 deaths are caused annually because of drugs side-effects which is 5 percent of the affected number of people in U.S. Drug-side-effects is a menace and it affects 6 percent of hospitalized patients. (Jahid & Ruan, 2013)

The application spectrum of studying side-effects spans concepts like combination therapies and drug repositioning. In complex cases like cancer where combination therapies are required, the drug-drug interaction has shown to cause more damage than good due to un-wanted adverse affects. Combination therapies administration for synergistic affect is failing because of lack of understanding of why side effects arise and how to limit them. There are multitude of potential reasons like unintended cascade of cross-talk between biomarkers in signaling pathways leading to unwanted reactions(Brechbiel, Miller-Moslin, & Adjei, 2014). Moreover, with the possibility of rapid estimation of in-silico drug side-effect through prediction modelling medium, the promising goal of drug-repositioning/drug-repurposing is also achievable (Yang & Agarwal, 2011). The idea of drug-repositioning is to maximize the use of already approved profiled drugs proven to be highly effective for some strata of patients could be recommended for a different and previously untested set of indications.

1.2 Related Work

Public health forums echo tension and highlights that there is an imminent requirement of research in this exciting era of modern medicine (Chee, Berlin, & Schatz, 2011). The concept of precision medicine is to find most suitable highest efficacy therapies for patients with minimal side-effects. As field of translational bioinformatics is rapidly advancing, it is possible to perform novel insilico drug-discoveries (Butte, 2008) Side-effects prediction of drugs is a critical and crucial step towards personalized medicine approach. Previously scientific studies for drug-side effect

prediction have used machine learning approaches (W. Zhang, Liu, Luo, & Zhang, 2015)(T.-B. Ho et al., 2016).

In this section, computational approaches which have been used for identifying drug side-effects are reviewed. Following summaries of previous approaches spans from cluster analysis, supervised deep learning strategy, factor analysis, causality analysis, network analysis and genome wide association studies (GWAS),enrichment analysis for result validations and data-mining approach. The data sources used to study side-effect in different studies include chemogenomic data of both drugs and drug-targets.

One recent advancement is DrugClust tool (Dimitri & Lió, 2017). It is an R package and uses machine learning to predict side-effects. There are two main steps in the analysis pipeline, cluster analysis followed by enrichment analysis. The data analysis pipeline first groups the drugs on the basis of similar features. Bayesian priors are assumed while doing this cluster analysis. As a second step enrichment analysis is performed for the clusters to extract a more biological interpretation of the clusters formed. The pathway enrichment analysis helps to investigate the interaction between drug clusters with complementary profiles, complementary profiles means that drugs which interact with similar drug-targets and interact with similar biological pathways and cause similar side-effects. Rand Index is a metric which is used to find the statistical significance of the clusters. The prediction performance has been shown on various publicly available datasets.

Bresso et al used an integrative approach to explain drugs side-effects. The data was acquired from Drugbank and SIDER database. Clustering of the similar drugs is performed by combining the drug targets descriptors and drugs fingerprints. Comparison of two machine learning methods i-e decision trees and inductive-logic programming shows that the later outperformed both in performance and to further explicate the functional association in pathways of drug targets and drugs. (Bresso et al., 2013)

An interesting approach to consider side-effect as penalty scores for the drugs to rank the drugs was adopted by *Niu et al*. After randomly generating scores in simulation experiments the average scores were used to rank the drugs. Three different data sources were compiled together for the study i-e drug targets, chemical descriptors of drugs and the treatment indications of the drugs. Ensemble machine learning models were used to assign different weights to drugs on the basis of different side-effects associated with the drugs, there intended targets and treatment indications. (Niu & Zhang, 2017) Awarding scores is an idea associated with gaming industry, used in this project to elaborate significant linkups between drug-diseases associations, drug and drug-side effects commonly caused by the medications used for treatment and it can aid researchers in pharmaceutical companies to generate hypotheses for drug discovery.

A link between pharmacogenomics and side-effects has been shown by isolating 244 pharmacogenes which are associated with side-effects of 176 drugs from PharmGKB database. 28 genes are identified by FDA which are associated with risk of side-effects. (Zhou et al., 2015) Another novel deep learning methodology for genome wide association studies (GWAS) to exploit the pharmacogenomic data and phenotypic

response in patients was conducted by *Liang et al.* This supervised deep learning methodology uses single nucleotide polymorphisms (SNPs), pharmacokinetic data and side-effects data. This model in particular classifies single nucleotide polymorphism (SNP) with adverse reactions. This model makes use of stochastic networks that rely on markov chains as step functions. This strategy outperformed baseline models like lasso regression and k-Nearest neighbour method. (Liang, Huang, Zeng, & Zhang, 2016)

Causality analysis model based on structure learning (CASTLE) tool uses both chemical and biological properties of drugs to determine molecular predictors of side-effects. Prediction performance was evaluated on 12 organ-specific ADRs on 830 drugs data. The analysis pipeline has three steps involved feature extraction, classification of ADRs using Support vector machines (SVM), enrichment analysis was performed for validation and compared with OMIM database results. Although the prediction performance was promising but there was only partial validation from enrichment analysis with OMIM database stands for mandelian inheritance traits in man and contains information related to mendelian disorders and over 15,000 genes. (Liu et al., 2014)

Another systematic prediction of ADRs was performed using factor analysis performed on 832 drugs. The data set is composed of 173 pathways and 1385 side-effects, 30 highly correlated sets of drugs, pathways and ADRs were identified as a result. The method used was canonical correlation analysis (CCA) which captures the co-variance between the predictive features and models the variation within the features as noise. This study is significant as the correlation between ADRs is significant from biological standpoint and this strategy lays the foundations for latent low dimentional representation of the chemogenomic data sources which could lead to better inferences and better understanding of the mechanisms that cause drug side-effects to appear in patients. (Zheng et al., 2014).

A large scale network analysis performed on side-effects and biological pathways required pathway data, target data and phenotypic data. Corelations between side-effects and associations between the biological pathways was the outcome of this study. (Scheiber et al., 2009). Data-mining approach has also been used to assist in pharmacovogilance as well.(Hauben, Horn, & Reich, 2007). Pharmacovigilance is the feedback drug surveillance from consumers to study and record the reported side-effects caused by medicines.

1.3 Rationale and motivation

The seriousness of this medical issue is reiterated by statistical figures pointing towards the fact that patients ailment is prolonged due to ADRs. Studies indicate that there is an average two day increase in hospitalization due to ADRs (Eichelbaum, Ingelman-Sundberg, & Evans, 2006). The risk of hospitalization due to ADRs in even higher in elderly patients aged greater than 75 years. (Ruiter et al., 2012). This increase in hospitalization time means more resources consumption by one patient overall, when the same healthcare budget could be utilized by other crucial patients in need. Another alarming fact is that drug-related mortality rate in U.S

alone is 4th, after cancer and heart problems.(Leone et al., 2008). The repercussions for investing in solving this complex medicine problem can have epic impact on "community medicine" and resource reservation in terms of both time and money.

Side-effects are a result of different off-target drug reactions. The mechanism of action of drugs is also influenced by the genomic heterogeneity of individuals and influencing chemical properties due to altering micro-environment in cellular compartments. Hence, side-effects (clinical phenotype) that arise in patients can assumed to be a manifestation of complex interaction of genomic features, chemical descriptors of drugs, drug-targets associated indications (Schuster et al., 2008). Finding out these underlying mechanisms of drugs side-effect that arise due to this complex interaction is the motivation behind this study.

1.4 Problem statement, challenge and research objectives

It is essential to establish an analysis pipeline to computationally predict drug side-effects from multiple diverse sources. The challenge in this study is that direct genetic information from the patients not accessible from public data repositories. Therefore, a fundamental assumption in this study is that drug indications are an approximation of missing genetic information from the patients. The basic research question this thesis aims to solve is if the chemical discriptors called as fingerprints of the drugs and drug indications are an insightful information source for the drug common side-effects reported with the drugs. The focus of the thesis is to classify ADRs associated with drug indications and chemical descriptors. ADRs caused due to other mechanisms are beyond the scope of this thesis. The research questions can be enlisted as following

- If the data sources selected are insightful for drug side-effect prediction or not. If so then which data source is more insightful information source for the classification of a particular side-effect.
- Which machine learning method can predict the side-effects using the data sources i-e best predictive model selection and assessment.
- How can side-effects be grouped together on the basis of there informative data source.

The data sources used to predict the side-effects are the known drug-disease associations and fingerprints/ chemical descriptors of the drugs. Ten different but common side-effects which were used for this study are namely Headache, Dizziness, Weakness, Abdominal Pain, Nausea, Chronic Fatigue, Diarrhea, Rashes, Dermatitis, Vomiting. These side-effects with highest variance were selected for this study. Precisely, data from therapeutic indications of drugs along with their chemical properties (chemical descriptors) are used to predict clinical phenotypes (side-effect) of drugs.

1.5 Structure and organization of the thesis

After presenting relevant background on the ADRs, Chapter 1 presented the challenges, basic assumptions and the research questions that this study aims to solve. Chapter 2 presents computational methods and data sources information used to predict side-effects, the data analysis pipeline and the cross-validation scheme. we learn multiple classifiers for the same data set with different views to model single ADRs, Chapter 3 presents the results section and Chapter 4 concludes the thesis with discussion, analysis and suggests directions for further research. Finally, we interpret prediction results and summarize all works in this thesis.

2 Research material and methods

The data sources used to predict the side-effects are the known drug-disease associations and fingerprints/ chemical descriptors of the drugs. Ten different but common side-effects which were used for this study are namely Headache, Dizziness, Weakness, Abdominal Pain, Nausea, Chronic Fatigue, Diarrhea, Rashes, Dermatitis, Vomiting. These side-effects with highest variance were selected for this study. Precisely, data from therapeutic indications of drugs along with their chemical properties (chemical descriptors) are used to predict clinical phenotypes (side-effect) of drugs.

2.1 Data acquisition and dataset construction

Data was collected from public data repositories CHEMBL and SIDER. These databases are freely accessible bioinformatics resource of information. ChEMBL is a bioactivity database which contains information that ranges from drug indications, Targets,e-t-c (Gaulton et al., 2011). SIDER is a public database that contains side-effects of drugs which include side effect frequency, drug side effect classifications and drug–target relations (Kuhn, Letunic, Jensen, & Bork, 2015).

2.1.1 Drug indications data

Drug-Disease relationship is known as drug indications. This drug-disease connectivity map data is retrieved from CHEMBL 22 database with MeSH ids (Medical Subject Headings (MeSH) vocabulary).

2.1.2 Drug descriptors and targets

The 2D fingerprints represent the structural properties of drugs like number of bonds and atoms (non-H atoms and rotatable bonds), functional groups, C-chains, Ring structure and size. MACCS fingerprints are one of the conventional examples of 2D fingerprints that represent drugs with a set of 166 fragments. These descriptors are used for investigating bio-activity of drugs.

2.1.3 Case study: Dataset statistics and Features description

Following dataset statistics summarize the constructed dataset for analysis. 2634 unique Drugs (CHEMBL compound Ids) were retrieved from CHEMBL along with 412 drug-target ids (tids), ATC codes. ID Cross-references with other databases namely STITCH database, PubChem database, Drugbank database are also retrieved using AWK and bash scripting.

For the number of features, there are 725 mesh ids (drug indications) and 166 ”MACCS” fingerprints calculated from the SMARTS codes retrieved from the CHEMBL 22 database and calculated using rcdk R package.

From the 1434 marketed drugs available, 2966 side-effects were retrieved from SIDER 4 with Medra Ids. which is a quantitative measurement recorded as side-effect frequency from 30 patients with recorded medical history. Overall side effect data

has in total 97.5 percent data as 0, .1 percent of data is 1 and 2.3 percent of the data matrix contains continuous values within 0 and 1. However, the final dataset that was used after preprocessing were complete cases with available side-effects is summarized in Table 1.

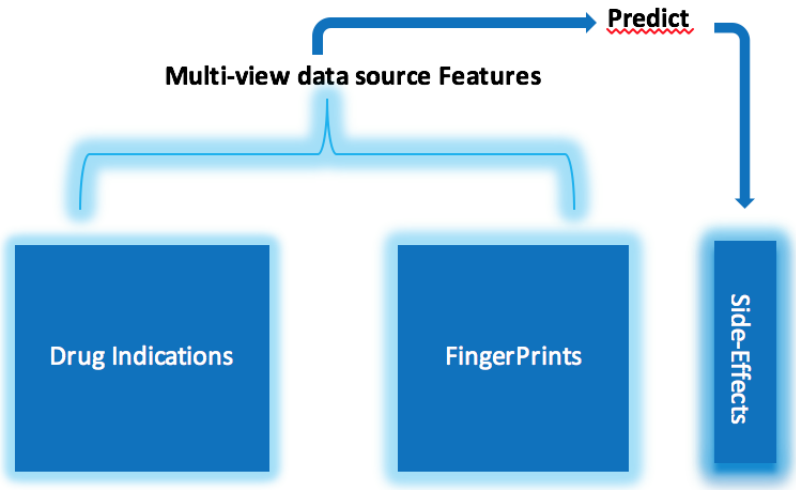

Figure 1: Features and target

Table 1: Summary of final dataset after preprocessing

No. of Drugs	Drug indications	fingerprints	No. of sideeffects
667	725	166	10

Drug indications predictors with zero standard deviation in column were removed, this reduced the dimensionality to 652 from 725. 10 side-effects which showed maximum variance across the drugs were selected for this prediction task and analysis. The predictors are all binary data but the target variable is a continuous data which ranges between 0 and 1. I have converted the quantitative data as the observations greater than 0.5 to 1 and lesser than 0.5 as 0. The sparsity of the final dataset used in the analysis has is summarized in Table 2.

Table 2: Sparsity of pre-processed data sources used for prediction performance

Drug indications	fingerprints	No. of side-effects
99.01%	70.81%	95.47%

2.2 Predictive Modelling and Experimental configuration

For the formulation of predictive models, the input to machine learning algorithms are feature vectors composed of drug indications and chemical fingerprints of drugs. Side-effect prediction task was effectively modelled as a binary classification problem where each drug was considered to either cause a particular side-effect (labelled 1) or not (labelled 0). As the side-effect data was in quantitative measure, a data preprocessing strategy is adopted, weak signals (<0.5) converted to 0 and stronger signals (>0.5) converted to 1. For biological drug-indications data and chemical descriptors (fingerprints) data, a total of 891 combined features were used for prediction of 10 general ADRs, a total of 7 predictive classifier linear and non-linear models were used.

In a 10 crossfold validation setup, with a 90/10 split only 10% of drugs are used to evaluate a classifier. The experimental setup ensures bootstrapping with 10 runs ensures each drug is tested. Balanced training set was ensured by interleaving the two classes in equal ratio as illustrated in the pipeline with 10-fold cross validation setup used in the present study.

Class imbalance problem arise after converting the weak signals to 0 and strong signals to 1. The negative class cases were $13:1$. The majority class was divided into twenty different chunks. Each majority class chunk becomes roughly equivalent of the much smaller positive class of interest. An approximate twenty different datasets were used as training samples for each of the ten bootstrap runs to evaluate the prediction performance of the classifier. One dataset was further sliced into 90:10 ratio of training set and test set to ensure unbiased prediction performance and to avoid over-fitting. Furthermore, a ten cross internal cross-validation was performed for the hyper-parameter optimization for different models. Collinearity was not explicitly removed as most models handle collinearity in features except Linear Discriminant Analysis. Finally, after all the 20 datasets for all bootstrap runs of the experiment. An average of the accuracy metric is evaluated. This is a robust analysis pipeline to evaluate the prediction performance of the classifiers. The analysis pipeline is also depicted in the fig 2.

A performance evaluation comparison for 7 different classifiers is the summarized in results section. Random assignment of classes to drugs is used as a baseline. The models include Naive Beyes, Linear Discriminatory Analysis, Support Vector machines, Logistic regression with lasso setting, Neural networks, Random Forest, Bayesian Efficient Multi-view method. For the comparison of machine learning models, three different kernels were implemented in SVM and BEMKL as jaccard, radial and linear kernels.

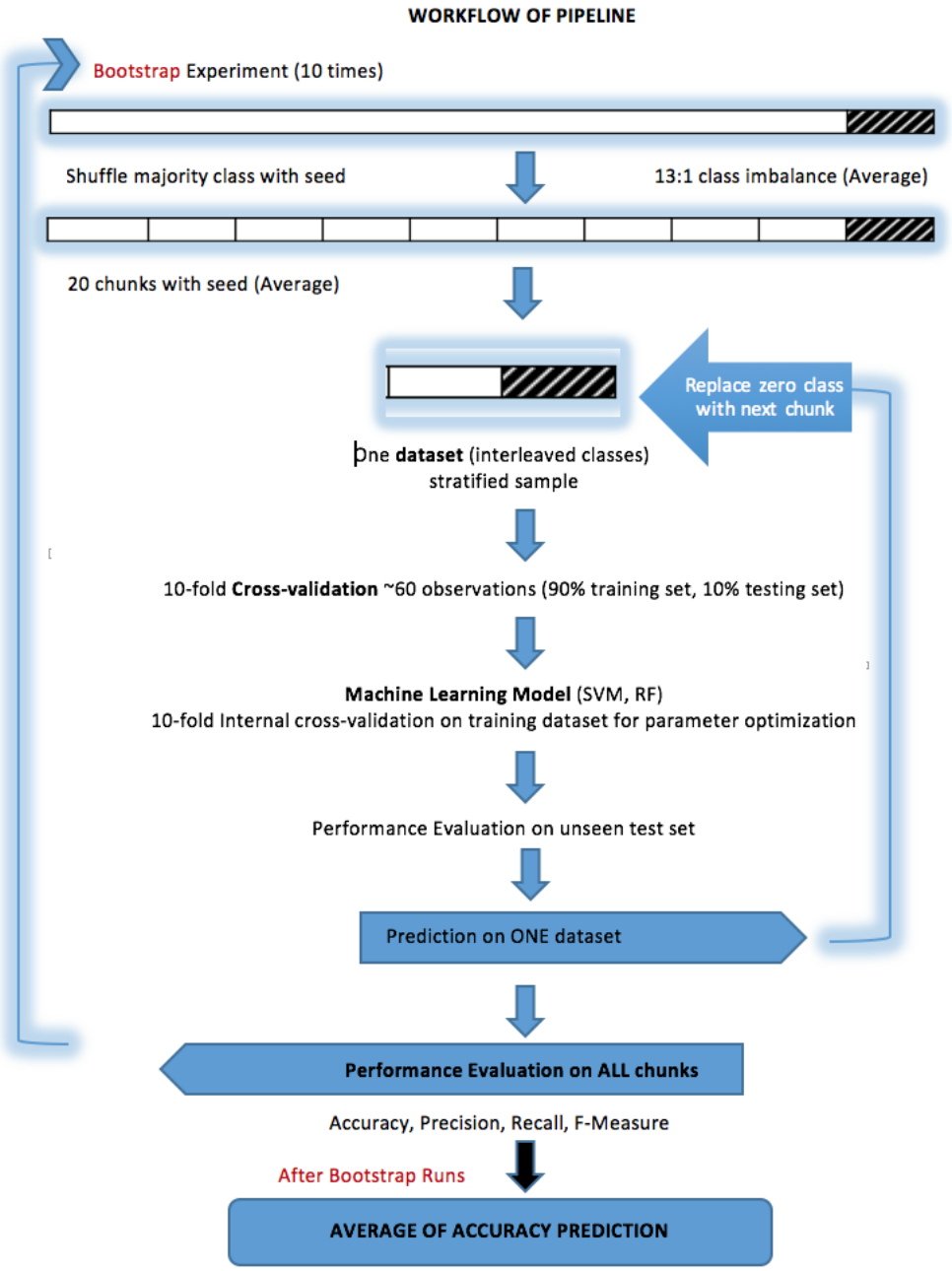

Figure 2: **Experimental Configuration**: cross validation scheme

2.3 Methods

This section presents a brief summary of the methods used in the analysis pipeline.

2.3.1 Naive Bayes

It is a supervised machine learning algorithm. Based on bayes theorem, with a basic assumption that the features are independent of one another. The covariance in features is not considered, hence the dimensionality of features does not affect the performance as each feature is considered to have independent distribution. (H. Zhang, 2004)

$$P(y|x_1, ...x_n) = P(y)\frac{P(x_1, ...x_n|y)}{P(x_1, ...x_n)} \tag{1}$$

y represents the side-effect class $P(y = 0/1)$ and x represents the features including either drug indications or drug fingerprints or both in case of combined views. The assumption is

$$P(x_i|y, x_1, ..., x_{i-1}, x_{i+1}, ..., x_n) = P(x_i|y) \tag{2}$$

For all the features represented by i,

$$P(y|x_1, ...x_n) = \frac{P(y)\ \prod_{i=1}^{n} P(x_i)|y)}{P(x_1, ..., x_n)} \tag{3}$$

As $P(x_1, ...x_n)$ is a constant, the classification rule is used

$$P(y|x_1, ...x_n) \propto P(y)\ \prod_{i=1}^{n} P(x_i)|y) \tag{4}$$

$$\hat{y} = \ arg\ max_y P(y)\ \prod_{i=1}^{n} P(x_i)|y) \tag{5}$$

Maximum A Posteriori (MAP) estimates the side-effect class y. *P(y)* is the prior probability of class that side-effect belongs to in the training set.

Despite the simple assumption, Naive bayes classifiers are used in practical applications. Parameter estimation requires less amount of training data. Naive bayes is a fast learner in comparison to advanced methods.

2.3.2 Linear Discriminant Analysis (LDA)

Linear Discriminant Analysis is a machine learning approach that finds linear combination of features that distinguishes two or more classes. The fundamental assumption of LDA is that independent variables are normally distributed. LDA models the difference between the two classes. The classification rule depends on the means of all features (drug indications data or fingerprints data), so result is affected by outliers in the data.

Mean and covariance are unknown and are estimated from training data. consider x drug sample observations from training set and for each sample y as the side-effect known class either 1 or 0. In a binary class prediction problem, LDA assumes the conditional probability density functions $P(x|y = 0)$ and $P(x|y = 1)$ are normally distributed with mean and covariance variables $(\mu_o \Sigma_o)$ and $(\mu_1 \Sigma_1)$. Additional

assumption by LDA is the identical class covariance for both positive class and negative class, called as homoscedasticity. ($\Sigma_0 = \Sigma_1 = \Sigma$) (Martínez & Kak, 2001).

2.3.3 Linear Regression

For linear regression, consider $\chi \epsilon \mathbb{R}^{n*p}$, the matrix of *drug indications* or *fingerprints* data and $y \epsilon \mathbb{R}^{n*1}$, the vector of side-effect. n denotes the independent and identically distributed number of observations (667 drug samples) and p represents the number of features. (chemical descriptors/drug indications).

Linear regression assumes that side-effect y is a resultant combination of unknown weight vector $w \epsilon \mathbb{R}^{1*p}$.

$$y = \chi \mathbf{w}^T + \epsilon \tag{6}$$

Noise ϵ exists in χ (error term in features space). Learning the weights $\mathbf{w}$ in feature matrix is the main goal in machine learning. With $\mathbf{w}$ learned, it is possible to narrow down important feature selection. Significant features provide significant biological insights in our data. Moreover, classification of side-effect is also possible. In "small n, large p" problems over-fitting on training data should be avoided to make improved inference as well. A workaround is to put penalty on weights and reduce the size of $\mathbf{w}$. the penalty system is called regularizing the cost function. we used lasso setting which means alpha was chosen as 1.

However, in order to predict discrete classes, logistic regression is used. Binary logistic regression for side-effect prediction is a simplistic classification problem to learn a function of the form in order to model dichotomous class.

$$P(y = 1|x) = h_\theta(x) = \frac{1}{1 + exp(-\theta^T x)} \equiv \sigma(\theta^T x) \tag{7}$$

$$P(y = 0|x) = 1 - P(y = 1|x) = 1 - h_\theta(x) \tag{8}$$

The function $\sigma(z) \equiv \frac{1}{1 + exp(-z)}$ is called sigmoid/logistic function. It converts the value $\theta^T x$ in the range of 0 and 1. so that we can predict the probability of class of side-effect. The aim is to estimate the value of theta so that the probability $P(y = 1|x) = h_\theta(x)$ is high when the x (whether drug indications / fingerprints) belong to class 1 of side-effect and vice-versa. In summary, if the probability of class 1 i-e $P(y = 1|x) > P(y = 0|x)$ then the side-effect class is predicted as class 1 and opposite otherwise (Cox, 1958).

2.3.4 Support Vector Machines (SVM)

Support vector machines is a maximum margin classifier. SVM can be applied for both linear class boundries and non-linear class boundaries, based on the type of kernel function invoked in the classifier. In a binary class setting, SVM is widely used for the classic binary classification problems because it has several advantages. SVM is a popular technique that efficiently handles large training sets. SVM is a discriminative classifier, the algorithm finds optimal hyperplane that separates

the positive instances from negative class instances and assigns maximal distance between the two classes. This is the case of the linear classification problem, however, in addition, kernel methods are available which transform non-linear space into linear ones for non-linear classification. kernels defines a similarity measure between two data points of drug observations and relative positions of the inputs in the feature space (drug indications / fingerprints).

Following equation represent the decision function.

$$f(x) = \mathbf{w}^T \kappa + b \tag{9}$$

$\mathbf{w}$ represents weights which are supposed to be learned for the samples, κ is the kernel (similarity measure in samples) and b is the error term. Default parameters were used for radial and jaccard kernel (James, Witten, Hastie, & Tibshirani, 2013).

Both linear and non-linear kernel methods are used to attempt to capture the significant signals in the data. Kernel functions are widely used to sequester signals in the data. If there is non-linearity in data then non-linear methods outperform the linear methods and vice versa. An implementation of Support Vector Machines (SVM), available from *e1071* R package and *libsvm* were used in this study.

2.3.5 Neural Networks

Neural Networks design is an adaptation of cerebral cortex model in simpler scale. Intuitively, the architecture of neural networks can be envisioned in terms of layers of connected nodes. There is an activation function associated with the layers to trigger the successive layers activation. The input layer is the first layer which represents the features (drug indications / fingerprints) and output layer is the class of side-effect to be predicted. The actual processing to assign weights to features and learn the signals are done by the activation function associated with the hidden layers. Back propagation is a widely used standard rule to assign weights to the features. Backpropagation rule used delta function. However, there are limitations to the model as it is slower to train (Caudill, 1989). The architectural composition selected for a problem is an arbitrary choice for the modeller to motivate the structure based on the problem. The internal working is such that the activation of one layer determines the activation of the following layer. The analogy is assumed to be an approximation of the biological networks of the neurons. A pattern of activation from the input layer related to features causes activation of pattern of internal nodes in the hidden layers, this stepwise trigger of activation in successive hidden layers in the architecture adopted ultimately leads to the output layer and the classification of side-effects class $P(y = 0/1)$ is made.

As there are two internal layers. In the first internal layer there are 30 nodes which implies that there are "n" number of features number which are connections from each input node to the first hidden layer. Every neuron in 1st hidden layer is connected to the feature input neuron with weights associated with them and a bias value to complete the sigmoid activation function. The sigmoid activation function is equivalent to the number of features * 30 weights and 30 bias value. In a combined view if we have 891 features in total, then 891 *30 + 30 *10 + 10 *2 total weights

and 30 + 10 + 2 total biases. 27050 weights and 42 bias in total. Finding the right weights and bias is the main task in machine learning. The architecture adopted for Neural Networks had two-hidden layers with 30 nodes in first layer and 10 in second inner hidden layer. As neural network is a hit-and-trial method the rationale for the selection of this architecture is the range of number of features as square-root of the total number of features in case of random forest.

2.3.6 Random Forest

Random forest is a supervised ensemble learning method. Working principle of a ensemble learning method is that a combination of weak learner are combined to make strong learners. Random forest builds numerous decision trees during training. Decision tree uses tree structure as it beaks down the dataset into smaller subsets. Hence, random forest is made by using and combining many decision tree models. Random Forest can be used for both classification and regression. Random forest handles over-fitting problem and handles the missing values problems as well (T. K. Ho, 1995). The generalization error depends on individual trees strength and correlation between them (Breiman, 2001). Random forest are widely used state-of-the-art non-linear methods for prediction analysis. Based on random selection of samples and features, random forest makes variety of trees. Prediction of new samples are made by averaging the predictions of different trees. A decision tree is built using the whole dataset whereas in a random forest only a pre-set fraction of features are randomly chosen with a different feature at the root node in each decision tree and particular number of features are used for training and decision trees are build. A number of decision trees are built to result an outcome. Random forest at the end compiles the results of all the decision trees to predict the final outcome i-e classification of side-effect.

2.3.7 Bayesian efficient multi kernel learning (BEMKL)

Bayesian efficient multi-kernel method was designed to integrate multiple views of information sources in a systematic way. This systematic integration leads to higher prediction performance. BEMKL is a sophisticated machine learning method developed recently to handle joint modelling biological data and it has been used to show improved drug sensitivity prediction (Costello et al., 2014).

BEMKL is unification of four modelling strategies multiview learning, multi-tasking learning, kernelized regression and Bayesian inference. However, multi-tasking is not invoked in this research. Input data (Drug indications/fingerprints) is represented by kernels. BEMKL has two-steps, firstly dimensionality reduction step is performed. In case of "small n large p" problems, the number of features becomes equal to number of samples. In second step the output is estimated from the kernel weights. The kernelized regression controls overfitting in "small n large p" problems. BEMKL gives as additive advantage to capture non-linearity from different sources of data. Moreover, the dimensionality reduction means that the feature space is reduced to the number of samples.

Integration of different data sources is the in a single model is also a plus point in BEMKL. Data fusion causes more insight into the data sources and better prediction performance is ensured. This systematic integration is absent in previous less sophisticated models. BEMKL learns kernel weights to find similarity measures in drugs. kernel weights highlight the significance of each data source in order to predict the drug side-effects (Gonen, 2012). Furthermore, default settings for hyper-parameters were utilized.

2.4 Evaluation of measures for predictive models

A total of 7 machine learning models were generated for 10 general adverse drug reactions, which were evaluated using multiple statistical measures as accuracy, precision, recall and F-measure. Accuracy is the quality ratio of correctly identified instances both positive and negative (TP+TN/(TP+TN+FP+FN)). Precision (P) is the only a fraction of correctly recognized positives cases of all the predicted positives cases (P=TP/(TP+FP)). True positive rate (TPR=TP/(TP+FN)) represents correctly identified positives which is also called recall while false positive rate(FPR=FP/(FP+TN)) represents correctly identified negatives. F-measure is the mean of precision and recall. The performance for the 7 models for 10 general ADRs was calculated for the given datasets summarized in the Table 3. For statistical significance of the results, t-test is performed to find the p value. The t score is a ratio between the difference between two groups and the difference within the groups. large t-score implies that the groups are different and vice versa. p-value (probability that the results came by chance) range from 0 percent to 100 percent represented by decimals. The lower the p-values the better it is, indicating that results from data did not occur by chance. The p values are adjusted using bonferoni adjustment, which is the most stringent method to adjust the p values. Bonferoni correction os to adjust the p value for Type I error (false positive) (Weisstein, 2001).

3 Results and Discussion

Using the skeletal prototype of the data analysis pipeline and algorithms used for modelling which is outlined in aforementioned section, this study addresses the research questions of this thesis. One of the concrete deductions from the study is that both drug indications and chemical descriptors of drugs called fingerprints are indeed informative sources in the prediction of drug side-effects. This validation is the result of the comparison of predictive performance of machine learning methods, most of which are better than random classifiers in most of the drug side-effects prediction as illustrated in Table 3. Analysis of different side-effects is discussed below in this section. The prediction performance plots are in the supplementary section. For statistical significance the t-test comparison of machine learning methods is summarized in Table 4-13 of supplementary section.

3.0.1 Abdominal pain

In case of Abdominal pain, as indicated in Table 3 in supplementary section, 60.10 accuracy for BEMKL with jaccard kernel outperforms all the models with both single view of drug indications and combined views of fingerprints and drug indications. BEMKL with linear kernel along with neural networks are the second best with 57 percent prediction performance. Table 4 indicates the statistical significance for test results for abdominal pain. T-test suggests that if the p value is less than 0.5 then the tests are significant and indicates that the method with the lesser value has higher accuracy than the other.

3.0.2 Chronic fever

In case of Chronic fever, BEMKL with jaccard kernel outperformed the other classifiers with prediction accuracy of 58.5 followed by support vector machines with jaccard kernel with 55 percent accuracy measure. All other methods performed poorly with combined views as compared to single view drug indications data. Fingerprints data was consistently poor contributor to the prediction performance of all methods. NN (with fingerprints) performed even poorly than random classification which indicates that fingerprints other than MACCS can also be explored for there insightfulness.

3.0.3 Dizziness

In case of Dizziness, BEMKL (Jaccard kernel) outperformed all methods with 60 percent accuracy using only drug indications data. All the methods consistently performed better using only drug indications data as compared to other two data sets of combnined views and single view of fingerprints. All methods performed poorly with fingerprints data due to which the combined view was also less informative. Methods BEMKL (J), BEMKL (R), BEMKL (L) and SVM (L) showed poor performance using only fingerprints data. Dizziness is a psychological condition which can occur because of multiple factors. It is difficult to directly associate it with the chemical

structure of drugs which could be one probable reason for the poor performance for less predictive signal in the fingerprint data source.

3.0.4 Dermatitis

In case of the dermatitis as shown in fig 5 in results section, prediction performance of BEMKL (Jaccard kernel) and random Forest is almost similar with approximately 62 percent with both combined views of fingerprints data and drug indications. Overall most of the methods performed better using the fingerprints data. All the other used methods have less than 60 percent prediction accuracy as indicated in the accuracy table. Table 12 also indicates that the results both in case of BEMKL(J) and Random forest are statistically significant.

3.0.5 Diarrhea

In case of diarrhea, combined views are overall more insightful and better prediction has resulted from systematic combination of the datasets. All the methods have shown better prediction using the combined views. BEMKL (J) has outperformed all other methods using combined views with more than 65 percent prediction performance Table 3 in supplementary table. Three methods BEMKL(Radial kernel), neural networks and support vector machines using radial kernel (SVM(R)) performed approximately equally well after BEMKL(J), Poor accuracy measures resulted by using fingerprints data.

3.0.6 Headache

In case of headache, the prediction performance of most of the algorithms had higher prediction accuracy with drug indications data. BEMKL (J) outperformed all other methods using combined views with 71 prediction accuracy. However, the prediction accuracy of BEMKL(J) using only drug indications was also high, approximately 70 percent, but around 56 percent using only fingerprints data. The prediction performance of BEMKL (radial kernel) was also quite high around 65 percent using combined views. Fingerprints data was consistent poor contributor in performance accuracy for all the methods and assumed to be less insightful as well. Furthermore, As a single view experiment, excluding BEMKL for all other methods drug indications were more informative and more insightful as illustrated in Fig 4.

3.0.7 Rashes

In case of Rashes, all methods have given better prediction performance using combined views except SVM (radial kernel) and Naïve bayes. All methods performed poorly using only drug indications data except BEMKL (linear kernel) and SVM (linear kernel). Overall, Neural networks with approximately 62 percent prediction performance has outperformed all other methods using combined views closely followed by BEMKL (J) using combined views.

3.0.8 Nausea

In case of Nausea, all methods have performed consistently well using single view of drug indications. All methods have performed consistently poor with fingerprints data. All methods have performed almost equally only in case of SVM (Jaccard kernel) using all the three data sources separately with approximately 52 percent prediction performance. SVM(L), SVM(R), Naïve bayes have performed worse than random classifier using only fingerprint data. The prediction performance of BEMKL (jaccard kernel, radial kernel and linear kernel) and random forest show prediction performance of 50 percent using only fingerprint data. Neural network show worst performance than random forest using fingerprint data with approximately 47 percent accuracy measure. Logistic regression show worst performance with all different datasets i-e combined views and single views separately. Fingerprints data is chemical descriptor data so it is difficult to associate it with psychological conditions.

3.0.9 Vomiting

In case of vomiting as shown in fig 14 in supplementary section, Overall, BEMKL (J) outperformed all other methods closely followed by BEMKL (R) using all datasets with prediction accuracy of around 58 percent. Although the prediction accuracy were different but BEMKL (R) and SVM (R) showed similarity in that they both had approximately same prediction accuracy for thee three datasets used for classification. Neural network showed odd behavior, performed particularly well with fingerprints data with approx 57 percent prediction prediction accuracy but performed worst in case of drug indications with approx 45 prediction accuracy.

3.0.10 Weakness

In case of weakness, BEMKL (Jaccard kernel) outperformed all other methods with combined views closely followed by BEMKL (radial kernel). In single view case, all methods performed better using drug indications data. In neural networks, SVM (L), SVM (R), LDA performed poorer than random classifier using the combined views because the models can not systematically sequester the signal and noise separately from the data sources.

3.1 Conclusions and Findings

In ”small n, large p” problems, concatenating data sources is a useful modelling technique. Joint learning of signal across multiple data sources provides more statistical strength. An essential assumption is that the combination of multiple data sources yields more relevant information and extracts the structure of signal from the data, which is critical for this drug side-effect prediction task. In this study, 8/10 side-effects cases, BEMKL (using Jaccard kernel) when jointly modelling the combined sources of drug indications and fingerprints has outperformed all the other methods whereas in other two (Nausea and Dizziness), BEMKL (Jaccard kernel) while jointly modelling the two different data sources is quite close to the best

prediction performance. This implies that additional information is useful than single information sources. It is an established fact that combined views represent more information which suggests that they could be more insightful. (Kuang et al., 2014) as illustrated in prediction of side-effect : diarrhea fig 3. This is a proof of concept that systematically integrating data from multiple sources yields higher prediction accuracy. Moreover, bayesian methods with jaccard kernel suits the binary data.

One of the important conclusions drawn from this study is that non-linear methods like BEMKL, neural networks and random forest has consistently outperformed the linear methods to predict the side-effects. This indicates that non-linearity is observed in data. In most of the biological problems non-linearity is normally observed in the data, hence this is also a proof of concept which hold true in this study. The accuracy measures are summarized in the Table 3 in supplementary section. For statistical significance the t-test comparison of machine learning methods is summarized in table 4-13 of supplementary section.

In order to find which data source is more insightful information source for the classification of a particular side-effect and which model can predict side-effects caused by drugs, datasets from drug indications and chemical descriptors are used for drug side-effect prediction independently as well. Another inference which can be drawn from this study is how well can side-effects be grouped together on the basis of there informative data source. This can be observed while analysing the prediction performance of machine learning models using single views of information sources, From the chosen set of side effects, side-effects can be grouped in two sets based on the source of information which is more insightful for the classification either drug indications and fingerprints. For cluster analysis, Dermatitis, Rashes and Vomiting are grouped together where fingerprints were more informative. A study conducted in 2004 used molecular modelling and QSAR (Quantitative stucture activity relationship) studies to study trypanosoma cruzi epimastigotes present in tsetse fly which causes itchiness. Principal component analysis (PCA) showed that the descriptors, N atom and attached 2 C atom from the dithiocarbamate group are significant indicator of the classification between the higher and lower trypanomicid activity. (Sanches, Taft, et al., 2004). so, literature review supports that drug fingerprints are an informative data source for skin diseases. Whereas Abdominal Pain, Chronic Fatigue, Headache, Dizziness, Nausea, Vomiting and Weakness are grouped where drug indications were more informative for side-effect classification. Most of these side-effects are co-occuring common biological symptoms associated with many diseases like food poisoning, Hepatitis, alimentary canal cancers, malaria, cholera (Friedman et al., 2008), (Beare, Taylor, Harding, Lewallen, & Molyneux, 2006), (Kuna & Gajewski, 2017) e-t-c.

In summary the findings of this study can be outlined as follows

- Both drug indications and fingerprints are insightful data sources for drug side-effects because machine learning models perform better than the random classifications.
- Bayesian efficient multi kernel learning method with jaccard kernel BEMKL (J) outperformed in joint modeling drug indications and fingerprints data.

- Overall non-linear methods like BEMKL, Neural network and neural networks have shown better prediction performance.
- In individual views, Abdominal pain, Chronic fatigue, dizziness, diarrhea, Nausea, Vomiting, Weakness can be grouped together as methods perform better using drug indications, side-effects like rashes, dermatitis are better predicted using fingerprints data

Prediction performance for three side-effects is shown in the results section Diarrhea, Headache, Dermatitis. Fig 3 is a presentation of diarrhea results highlighting that joint modelling by BEMKL using jaccard kernel outperformed all other methods. Fig 4 represents presentation of Headache results highlighting that in single views, out of the two data sources used, drug indications is more informative and insightful data source. Fig 5 shows dermatitis results highlighting that in skin diseases related side-effects, fingerprints data is more insightful in side-effect prediction. Statistical significance machine learning methods is summarized in Table 4-13 of supplementary section.

3.2 Limitations and Potential Future Directions

A key limitation is the non-availability of genomic information of patients. Performance enhancement is expected with the additional information of actual genomic features of patients. If we are able to predict the common side effect without genomic information of the patients, then we can compare prediction performance with the addition of the actual genomic information and quantify the impact of genomic influence. The difference in the side-effect prediction accuracy measure would reflect how much genomic features contribute towards adverse effects of drugs.

A wide range of hypothesis in the theme of drug-repositioning can also be generated further with the extension of this research work. However, a potential improvement in analysis strategy is to take the covariance in all the target side-effects into account. It is pertinent to apply advanced machine learning methods to study multiple sets of data together. This thesis work lays the groundwork to adopt multi-task strategy and use Bayesian component based kernelized matrix factorization with flat and wide priors to explicate the structure in the data. Flat and wide priors means that the model learns on its own which is used in absence of any expert opinion and prior knowledge about the expectation of the results. In order to better solve this problem, the chemical information of drugs on cell-lines / detailed patient's molecular responses along with the phenotypic properties needs to be combined through a systematic method in order to learn the complex inter-dependencies, to predict the side-effects set.

To highlight the promising future directions, latent feature approach to use multi-task learning models should be investigated for side-effect prediction. Hyper-parameter tweaking and optimisation under bayesian setting to can improve the prediction results. Moreover, regression approaches like MVLR will also be explored.(Ammad-ud din, Khan, Wennerberg, & Aittokallio, 2017). This strategy can then assist in improving understanding of the molecular mechanisms of drugs.

This work aims to aid investigative exploratory analysis for safe drug's administration using multiple machine learning classifiers. Intuitively in retrospect this research approach is synonymous with the concept of reverse engineering i-e to indirectly quantify the impact of genomic features. The drug indications (drug-disease connection) and chemical descriptors (fingerprints) data is used as input to predict the general side-effects. A comparison with the prediction performance of different models after the addition of genomic features for the same patients would illustrate the entire picture uncovering meaningful biological insights leading to comprehensive understanding of drug-responses in patients. Ultimately, this study would have downstream implications on personalized medicine, which could be significantly boosted with the proper identification of suitable drugs for the individual patients which requires continuous work in this research domain. Thus, this study contributes to many ongoing projects for machine learning based side effect classification. This analysis pipeline may serve as a baseline platform to detect the side effects of preclinical drugs to facilitate and extend further research.

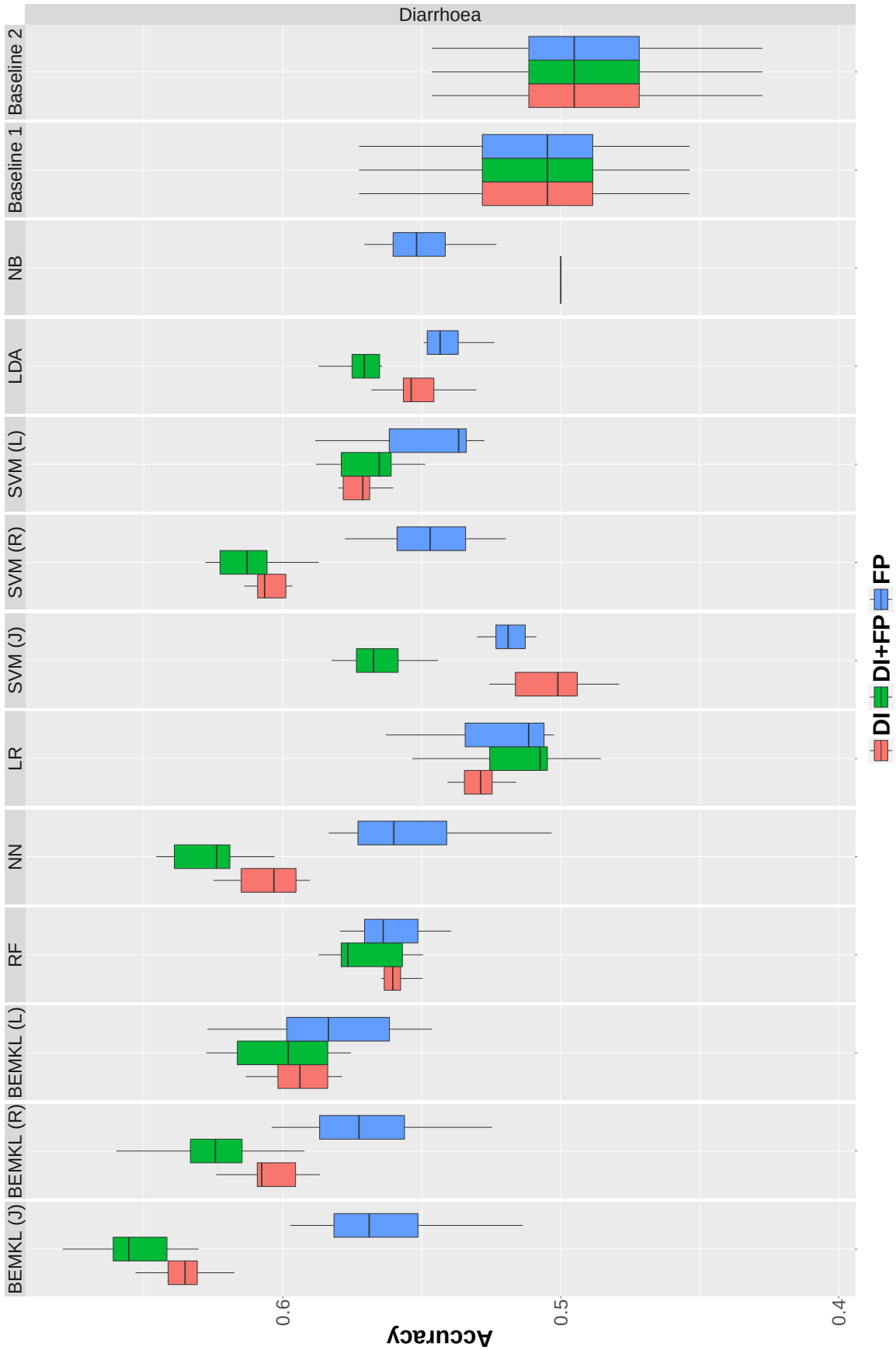

Figure 3: Performance measure for Side-effect : Diarrhoea

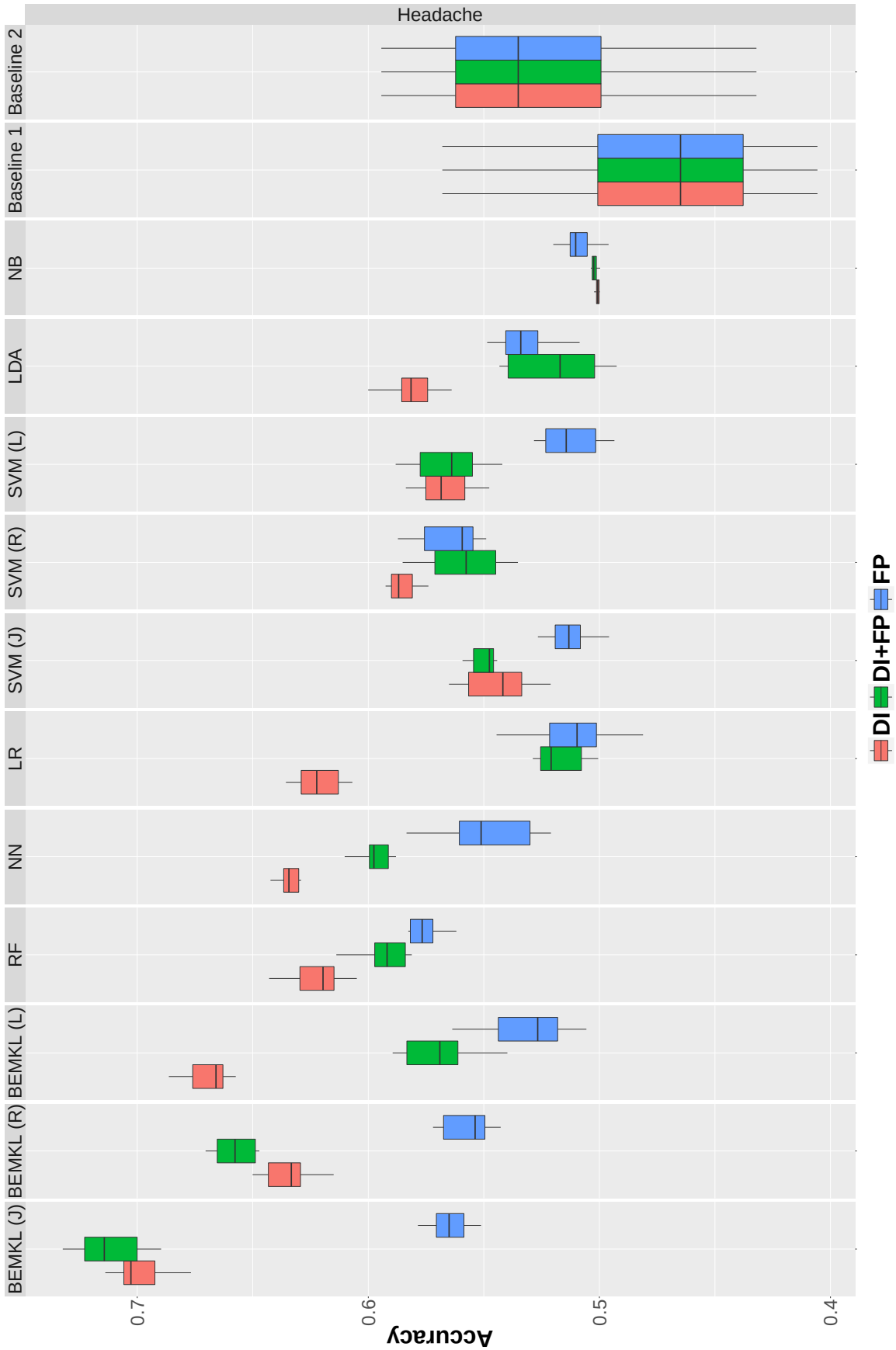

Figure 4: Performance measure for Side-effect : Headache

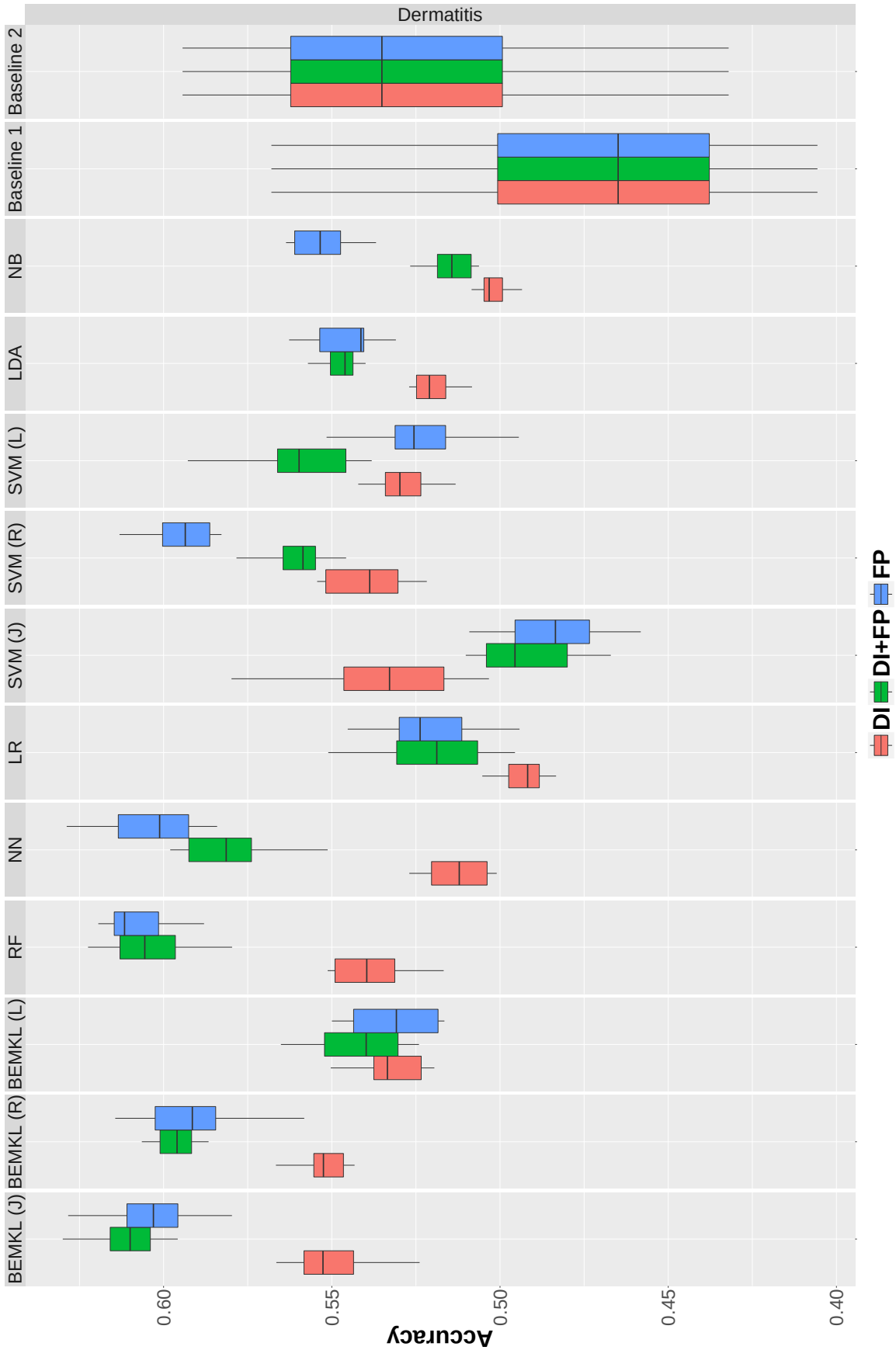

Figure 5: Performance measure for Side-effect : Dermatitis

References

Ammad-ud din, M., Khan, S. A., Wennerberg, K., & Aittokallio, T. (2017). Systematic identification of feature combinations for predicting drug response with bayesian multi-view multi-task linear regression. *Bioinformatics*, *33*(14), i359–i368.

Beare, N. A., Taylor, T. E., Harding, S. P., Lewallen, S., & Molyneux, M. E. (2006). Malarial retinopathy: a newly established diagnostic sign in severe malaria. *The American journal of tropical medicine and hygiene*, *75*(5), 790–797.

Bloomquist, L. (n.d.). How pharmaceuticals came to be the 4th leading cause of death in america.

Bracken, M. B. (2009). Why animal studies are often poor predictors of human reactions to exposure. *Journal of the royal society of medicine*, *102*(3), 120–122.

Brechbiel, J., Miller-Moslin, K., & Adjei, A. A. (2014). Crosstalk between hedgehog and other signaling pathways as a basis for combination therapies in cancer. *Cancer treatment reviews*, *40*(6), 750–759.

Breiman, L. (2001). Random forests. *Machine learning*, *45*(1), 5–32.

Bresso, E., Grisoni, R., Marchetti, G., Karaboga, A. S., Souchet, M., Devignes, M.-D., & Smaïl-Tabbone, M. (2013). Integrative relational machine-learning for understanding drug side-effect profiles. *BMC bioinformatics*, *14*(1), 207.

Butte, A. J. (2008). Translational bioinformatics: coming of age. *Journal of the American Medical Informatics Association*, *15*(6), 709–714.

Caudill, M. (1989). Neural nets primer, part vi. *AI Expert*, *4*(2), 61–67.

Chee, B. W., Berlin, R., & Schatz, B. (2011). Predicting adverse drug events from personal health messages. In *Amia annual symposium proceedings* (Vol. 2011, p. 217).

Chiang, A. P., & Butte, A. J. (2009). Data-driven methods to discover molecular determinants of serious adverse drug events. *Clinical Pharmacology & Therapeutics*, *85*(3), 259–268.

Costello, J. C., Heiser, L. M., Georgii, E., Gönen, M., Menden, M. P., Wang, N. J., ... others (2014). A community effort to assess and improve drug sensitivity prediction algorithms. *Nature biotechnology*, *32*(12), 1202–1212.

Cox, D. R. (1958). The regression analysis of binary sequences. *Journal of the Royal Statistical Society. Series B (Methodological)*, 215–242.

DiMasi, J. A. (2002). The value of improving the productivity of the drug development process. *Pharmacoeconomics*, *20*(3), 1–10.

Dimitri, G. M., & Lió, P. (2017). Drugclust: A machine learning approach for drugs side effects prediction. *Computational Biology and Chemistry*, *68*, 204–210.

Eichelbaum, M., Ingelman-Sundberg, M., & Evans, W. E. (2006). Pharmacogenomics and individualized drug therapy. *Annu. Rev. Med.*, *57*, 119–137.

Friedman, S., Blumberg, R. S., Kasper, D., Braunwald, E., Fauci, A., Hauser, S., ... Jameson, J. (2008). Harrison's principles of internal medicine.

Gaulton, A., Bellis, L. J., Bento, A. P., Chambers, J., Davies, M., Hersey, A., ... others (2011). Chembl: a large-scale bioactivity database for drug discovery.

Nucleic acids research, *40*(D1), D1100–D1107.

Gonen, M. (2012). Bayesian efficient multiple kernel learning. *arXiv preprint arXiv:1206.6465*.

Hauben, M., Horn, S., & Reich, L. (2007). Potential use of data-mining algorithms for the detection of 'surprise'adverse drug reactions. *Drug safety*, *30*(2), 143–155.

Ho, T.-B., Le, L., Thai, D. T., & Taewijit, S. (2016). Data-driven approach to detect and predict adverse drug reactions. *Current pharmaceutical design*, *22*(23), 3498–3526.

Ho, T. K. (1995). Random decision forests. In *Document analysis and recognition, 1995., proceedings of the third international conference on* (Vol. 1, pp. 278–282).

Huang, L.-C., Wu, X., & Chen, J. Y. (2011). Predicting adverse side effects of drugs. *BMC genomics*, *12*(5), S11.

Jahid, M. J., & Ruan, J. (2013). An ensemble approach for drug side effect prediction. In *Bioinformatics and biomedicine (bibm), 2013 ieee international conference on* (pp. 440–445).

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning* (Vol. 112). Springer.

Kola, I., & Landis, J. (2004). Can the pharmaceutical industry reduce attrition rates? *Nature reviews Drug discovery*, *3*(8), 711–716.

Kuang, Q., Wang, M., Li, R., Dong, Y., Li, Y., & Li, M. (2014). A systematic investigation of computation models for predicting adverse drug reactions (adrs). *PloS one*, *9*(9), e105889.

Kuhn, M., Letunic, I., Jensen, L. J., & Bork, P. (2015). The sider database of drugs and side effects. *Nucleic acids research*, *44*(D1), D1075–D1079.

Kuna, A., & Gajewski, M. (2017). Cholera—the new strike of an old foe. *International Maritime Health*, *68*(3), 163–167.

Leone, R., Sottosanti, L., Iorio, M. L., Santuccio, C., Conforti, A., Sabatini, V., ... Venegoni, M. (2008). Drug-related deaths. *Drug Safety*, *31*(8), 703–713.

Liang, Z., Huang, J. X., Zeng, X., & Zhang, G. (2016). Dl-adr: a novel deep learning model for classifying genomic variants into adverse drug reactions. *BMC medical genomics*, *9*(2), 48.

Liu, M., Cai, R., Hu, Y., Matheny, M. E., Sun, J., Hu, J., & Xu, H. (2014). Determining molecular predictors of adverse drug reactions with causality analysis based on structure learning. *Journal of the American Medical Informatics Association*, *21*(2), 245–251.

Ma, X. H., Wang, R., Xue, Y., Li, Z. R., Yang, S. Y., Wei, Y. Q., & Chen, Y. Z. (2008). Advances in machine learning prediction of toxicological properties and adverse drug reactions of pharmaceutical agents. *Current drug safety*, *3*(2), 100–114.

Martínez, A. M., & Kak, A. C. (2001). Pca versus lda. *IEEE transactions on pattern analysis and machine intelligence*, *23*(2), 228–233.

Niu, Y., & Zhang, W. (2017). Quantitative prediction of drug side effects based on drug-related features. *Interdisciplinary Sciences: Computational Life Sciences*, 1–11.

Ruiter, R., Visser, L. E., Rodenburg, E. M., Trifiro, G., Ziere, G., & Stricker, B. H. (2012). Adverse drug reaction-related hospitalizations in persons aged 55 years and over. *Drugs & aging*, *29*(3), 225–232.

Sanches, S. M., Taft, C. A., et al. (2004). A molecular modeling and qsar study of suppressors of the growth of trypanosoma cruzi epimastigotes. *Journal of Molecular Graphics and Modelling*, *23*(1), 89–97.

Scheiber, J., Jenkins, J. L., Sukuru, S. C. K., Bender, A., Mikhailov, D., Milik, M., ... others (2009). Mapping adverse drug reactions in chemical space. *Journal of medicinal chemistry*, *52*(9), 3103–3107.

Schuster, D., Laggner, C., & Langer, T. (2008). Why drugs fail–a study on side effects in new chemical entities. *Antitargets. Prediction and Prevention of Drug Side Effects*, 3–22.

Terstappen, G. C., & Reggiani, A. (2001). In silico research in drug discovery. *Trends in pharmacological sciences*, *22*(1), 23–26.

Weisstein, E. W. (2001). Student's t-distribution. *sigma*, *13*, 14.

Yang, L., & Agarwal, P. (2011). Systematic drug repositioning based on clinical side-effects. *PloS one*, *6*(12), e28025.

Zhang, H. (2004). The optimality of naive bayes. *AA*, *1*(2), 3.

Zhang, J. (2012). *Multi-task and multi-view learning for predicting adverse drug reactions* (Unpublished doctoral dissertation). University of Kansas.

Zhang, W., Liu, F., Luo, L., & Zhang, J. (2015). Predicting drug side effects by multi-label learning and ensemble learning. *BMC bioinformatics*, *16*(1), 365.

Zheng, H., Wang, H., Xu, H., Wu, Y., Zhao, Z., & Azuaje, F. (2014). Linking biochemical pathways and networks to adverse drug reactions. *IEEE transactions on nanobioscience*, *13*(2), 131–137.

Zhou, Z.-W., Chen, X.-W., Sneed, K. B., Yang, Y.-X., Zhang, X., He, Z.-X., ... Zhou, S.-F. (2015). Clinical association between pharmacogenomics and adverse drug reactions. *Drugs*, *75*(6), 589–631.

4 Supplementary Section

The supplementary section contains statistical measures for different side effcts with different machine learning models and T-tests for different side-effects.

Table 3: Statistic Measure for different sideeffects with different machine learning models.

Abdominal Pain				
Models	Accuracy	Precision	Recall	F-measure
BEMKL_J_DIFP	0.6001 ±0.016	0.01545	0.4304	0.02984
BEMKL_J_DI	0.6011 ±0.013	0.01390	0.3870	0.02684
BEMKL_J_FP	0.4831 ±0.015	0.01424	0.3609	0.02740
BEMKL_L_DIFP	0.4794 ±0.019	0.01765	0.4761	0.03404
BEMKL_L_DI	0.5724 ±0.009	0.01424	0.3312	0.02731
BEMKL_L_FP	0.4804 ±0.016	0.03110	0.5080	0.05862
BEMKL_R_DIFP	0.5482 ±0.023	0.01704	0.4652	0.03288
BEMKL_R_DI	0.5576 ±0.010	0.01375	0.3441	0.02643
BEMKL_R_FP	0.4855 ±0.017	0.01281	0.3413	0.02469
RF_DIFP	0.5094 ±0.013	0.51766	0.5159	0.51679
RF_DI	0.5320 ±0.011	0.47307	0.5506	0.50888
RF_FP	0.5015 ±0.013	0.51706	0.5087	0.51285
NN_DIFP	0.5330 ±0.020	0.01812	0.5043	0.03499
NN_DI	0.5729 ±0.012	0.01890	0.5261	0.03649
NN_FP	0.5361 ±0.015	0.02281	0.6348	0.04404
LR_DIFP	0.4945 ±0.008	0.80336	0.5087	0.62293
LR_DI	0.5107 ±0.007	0.92332	0.5148	0.66106
LR_FP	0.4963 ±0.009	0.79891	0.5103	0.62276

Continuation of Table 3				
Models	Accuracy	Precision	Recall	F-measure
SVM_J_DIFP	0.5050 ±0.012	0.50457	0.5131	0.50878
SVM_J_DI	0.5144 ±0.006	0.68661	0.5153	0.58872
SVM_J_FP	0.5012 ±0.015	0.52656	0.5097	0.51801
SVM_L_DIFP	0.5111 ±0.018	0.49107	0.5201	0.50516
SVM_L_DI	0.5369 ±0.008	0.53640	0.5489	0.54260
SVM_L_FP	0.4810 ±0.012	0.48444	0.4879	0.48618
SVM_R_DIFP	0.5242 ±0.010	0.49171	0.5392	0.51434
SVM_R_DI	0.5320 ±0.010	0.35220	0.5752	0.43688
SVM_R_FP	0.4985 ±0.016	0.54328	0.5076	0.52484
LDA_DIFP	0.5114 ±0.014	0.51307	0.5180	0.51552
LDA_DI	0.5445 ±0.010	0.55859	0.5514	0.55498
LDA_FP	0.4954 ±0.015	0.49628	0.5029	0.49958
NB_DIFP	0.5063 ±0.006	0.62130	0.5426	0.57930
NB_DI	0.5062 ±0.004	0.67864	0.5250	0.59204
NB_FP	0.5054 ±0.017	0.38715	0.4670	0.42336
RANDOM_P_DIFP	0.4976 ±0.069	0.52177	0.5007	0.51103
RANDOM_P_DI	0.4976 ±0.069	0.52177	0.5007	0.51103
RANDOM_P_FP	0.4976 ±0.069	0.52177	0.5007	0.51103
RANDOM_W_DIFP	0.5024 ±0.069	0.47823	0.5043	0.49091
RANDOM_W_DI	0.5024 ±0.069	0.47823	0.5043	0.49091
RANDOM_W_FP	0.5024 ±0.069	0.47823	0.5043	0.49091
Chronic Fever				
Models	Accuracy	Precision	Recall	F-measure
BEMKL_J_DIFP	0.5856 ±0.015	0.0188	0.5217	0.0363
BEMKL_J_DI	0.5679 ±0.012	0.0143	0.3935	0.0275
BEMKL_J_FP	0.5079 ±0.017	0.0214	0.5205	0.0412
BEMKL_L_DIFP	0.5105 ±0.009	0.0150	0.4109	0.0290
BEMKL_L_DI	0.5312 ±0.015	0.0125	0.3174	0.0241
BEMKL_L_FP	0.5169 ±0.016	0.0297	0.4274	0.0555
BEMKL_R_DIFP	0.5478 ±0.017	0.0190	0.5261	0.0366
BEMKL_R_DI	0.5666 ±0.015	0.0225	0.4844	0.0430
BEMKL_R_FP	0.5035 ±0.018	0.0174	0.4167	0.0333
RF_DIFP	0.5081 ±0.013	0.5332	0.5154	0.5241
RF_DI	0.5405 ±0.011	0.6812	0.5328	0.5979
RF_FP	0.5163 ±0.013	0.5113	0.5234	0.5173
NN_DIFP	0.4706 ±0.014	0.0139	0.3870	0.0269
NN_DI	0.5486 ±0.012	0.0164	0.4565	0.0318
NN_FP	0.4707 ±0.009	0.0142	0.3957	0.0275
LR_DIFP	0.5066 ±0.012	0.7601	0.5170	0.6154
LR_DI	0.5199 ±0.011	0.9176	0.5207	0.6644
LR_FP	0.5056 ±0.010	0.7429	0.5164	0.6093
SVM_J_DIFP	0.5525 ±0.020	0.5295	0.5544	0.5417
SVM_J_DI	0.5010 ±0.021	0.4459	0.5008	0.4717
SVM_J_FP	0.5251 ±0.011	0.4913	0.5238	0.5070
SVM_L_DIFP	0.5122 ±0.015	0.5201	0.5189	0.5195
SVM_L_DI	0.5573 ±0.009	0.6479	0.5508	0.5955
SVM_L_FP	0.5109 ±0.013	0.5035	0.5181	0.5107
SVM_R_DIFP	0.4702 ±0.017	0.5479	0.4824	0.5130
SVM_R_DI	0.4836 ±0.019	0.5492	0.4939	0.5201
SVM_R_FP	0.5097 ±0.011	0.4973	0.5209	0.5088
LDA_DIFP	0.4709 ±0.014	0.4743	0.4781	0.4762
LDA_DI	0.5456 ±0.010	0.6229	0.5415	0.5793
LDA_FP	0.5045 ±0.010	0.4940	0.5126	0.5031
NB_DIFP	0.5021 ±0.006	0.5760	0.5162	0.5445
NB_DI	0.5062 ±0.005	0.6201	0.5379	0.5761
NB_FP	0.5008 ±0.007	0.3659	0.4847	0.4170
RANDOM_P_DIFP	0.4958 ±0.066	0.5183	0.4985	0.5082
RANDOM_P_DI	0.4958 ±0.066	0.5183	0.4985	0.5082
RANDOM_P_FP	0.4958 ±0.066	0.5183	0.4985	0.5082
RANDOM_W_DIFP	0.5042 ±0.066	0.4817	0.5058	0.4935
RANDOM_W_DI	0.5042 ±0.066	0.4817	0.5058	0.4935
RANDOM_W_FP	0.5042 ±0.066	0.4817	0.5058	0.4935
Dizziness				
Models	Accuracy	Precision	Recall	F-measure
BEMKL_J_DIFP	0.5696 ±0.027	0.0147	0.4087	0.0283
BEMKL_J_DI	0.5969 ±0.023	0.0178	0.4957	0.0344
BEMKL_J_FP	0.4610 ±0.018	0.0085	0.1942	0.0162
BEMKL_L_DIFP	0.4452 ±0.011	0.0103	0.2746	0.0199
BEMKL_L_DI	0.5656 ±0.020	0.0233	0.6333	0.0449
BEMKL_L_FP	0.4343 ±0.015	0.0297	0.3458	0.0546
BEMKL_R_DIFP	0.5033 ±0.025	0.0109	0.2957	0.0211
BEMKL_R_DI	0.5481 ±0.021	0.0237	0.5418	0.0455
BEMKL_R_FP	0.4485 ±0.015	0.0078	0.1942	0.0150
RF_DIFP	0.4841 ±0.014	0.5134	0.4912	0.5020
RF_DI	0.5373 ±0.019	0.5768	0.5388	0.5571
RF_FP	0.4772 ±0.009	0.5066	0.4845	0.4953
NN_DIFP	0.5259 ±0.018	0.0167	0.4652	0.0323
NN_DI	0.5701 ±0.019	0.0205	0.5696	0.0395
NN_FP	0.4750 ±0.020	0.0117	0.3261	0.0226
LR_DIFP	0.4820 ±0.009	0.8318	0.4979	0.6229
LR_DI	0.5003 ±0.007	0.9315	0.5090	0.6583
LR_FP	0.4847 ±0.008	0.8221	0.4994	0.6213
SVM_J_DIFP	0.4801 ±0.010	0.5110	0.4892	0.4999
SVM_J_DI	0.5551 ±0.013	0.7031	0.5495	0.6169
SVM_J_FP	0.4883 ±0.020	0.5113	0.4952	0.5031
SVM_L_DIFP	0.5117 ±0.011	0.4996	0.5183	0.5088
SVM_L_DI	0.5371 ±0.021	0.5305	0.5405	0.5354
SVM_L_FP	0.4531 ±0.016	0.4458	0.4581	0.4518
SVM_R_DIFP	0.4996 ±0.016	0.5690	0.5054	0.5353
SVM_R_DI	0.4960 ±0.015	0.4394	0.5091	0.4717
SVM_R_FP	0.4879 ±0.010	0.5621	0.4983	0.5283

Continuation of Table 3				
Models	Accuracy	Precision	Recall	F-measure
LDA_DIFP	0.4806 ±0.014	0.4878	0.4878	0.4878
LDA_DI	0.5532 ±0.015	0.5191	0.5603	0.5389
LDA_FP	0.5042 ±0.013	0.4963	0.5117	0.5039
NB_DIFP	0.5019 ±0.007	0.6256	0.5023	0.5572
NB_DI	0.5033 ±0.004	0.6994	0.5140	0.5925
NB_FP	0.4900 ±0.012	0.6874	0.4766	0.5629
RANDOM_P_DIFP	0.4976 ±0.069	0.5218	0.5007	0.5110
RANDOM_P_DI	0.4976 ±0.069	0.5218	0.5007	0.5110
RANDOM_P_FP	0.4976 ±0.069	0.5218	0.5007	0.5110
RANDOM_W_DIFP	0.5024 ±0.069	0.4782	0.5043	0.4909
RANDOM_W_DI	0.5024 ±0.069	0.4782	0.5043	0.4909
RANDOM_W_FP	0.5024 ±0.069	0.4782	0.5043	0.4909
Headache				
Models	Accuracy	Precision	Recall	F-measure
BEMKL_J_DIFP	0.7117 ±0.014	0.019	0.6200	0.0377
BEMKL_J_DI	0.6988 ±0.011	0.018	0.5750	0.0355
BEMKL_J_FP	0.5645 ±0.008	0.017	0.4675	0.0330
BEMKL_L_DIFP	0.5689 ±0.016	0.019	0.5396	0.0367
BEMKL_L_DI	0.6671 ±0.013	0.027	0.7146	0.0514
BEMKL_L_FP	0.5322 ±0.019	0.043	0.5240	0.0803
BEMKL_R_DIFP	0.6556 ±0.014	0.020	0.6150	0.0386
BEMKL_R_DI	0.6351 ±0.011	0.020	0.5542	0.0379
BEMKL_R_FP	0.5568 ±0.010	0.016	0.4171	0.0303
RF_DIFP	0.5924 ±0.010	0.602	0.6056	0.6038
RF_DI	0.6222 ±0.012	0.571	0.6601	0.6124
RF_FP	0.5798 ±0.013	0.626	0.5882	0.6064
NN_DIFP	0.5992 ±0.017	0.020	0.6400	0.0390
NN_DI	0.6308 ±0.010	0.022	0.7000	0.0426
NN_FP	0.5485 ±0.020	0.012	0.3900	0.0238
LR_DIFP	0.5213 ±0.020	0.833	0.5336	0.6506
LR_DI	0.6179 ±0.018	0.848	0.6003	0.7028
LR_FP	0.5110 ±0.018	0.860	0.5258	0.6527
SVM_J_DIFP	0.5514 ±0.008	0.587	0.5634	0.5748
SVM_J_DI	0.5440 ±0.014	0.721	0.5466	0.6220
SVM_J_FP	0.5135 ±0.009	0.443	0.5312	0.4829
SVM_L_DIFP	0.5664 ±0.015	0.567	0.5808	0.5738
SVM_L_DI	0.5666 ±0.011	0.522	0.5978	0.5576
SVM_L_FP	0.5119 ±0.012	0.511	0.5265	0.5188
SVM_R_DIFP	0.5589 ±0.017	0.572	0.5796	0.5757
SVM_R_DI	0.5866 ±0.008	0.414	0.6713	0.5123
SVM_R_FP	0.5643 ±0.014	0.653	0.5715	0.6094
LDA_DIFP	0.5192 ±0.019	0.472	0.5367	0.5021
LDA_DI	0.5803 ±0.013	0.530	0.6094	0.5669
LDA_FP	0.5319 ±0.012	0.516	0.5467	0.5309
NB_DIFP	0.5022 ±0.001	0.878	0.5373	0.6665
NB_DI	0.5010 ±0.001	0.967	0.5192	0.6755
NB_FP	0.5094 ±0.007	0.843	0.5164	0.6404
RANDOM_P_DIFP	0.5215 ±0.054	0.545	0.5366	0.5407
RANDOM_P_DI	0.5215 ±0.054	0.545	0.5366	0.5407
RANDOM_P_FP	0.5215 ±0.054	0.545	0.5366	0.5407
RANDOM_W_DIFP	0.4785 ±0.054	0.455	0.4933	0.4734
RANDOM_W_DI	0.4785 ±0.054	0.455	0.4933	0.4734
RANDOM_W_FP	0.4785 ±0.054	0.455	0.4933	0.4734
Nausea				
Models	Accuracy	Precision	Recall	F-measure
bemkl_J_DIFP	0.5746 ±0.018	0.0187	0.5900	0.0362
bemkl_J_DI	0.5850 ±0.016	0.0158	0.4975	0.0306
bemkl_J_FP	0.4954 ±0.018	0.0106	0.3083	0.0204
bemkl_L_DIFP	0.5025 ±0.016	0.0169	0.4950	0.0327
bemkl_L_DI	0.5402 ±0.014	0.0194	0.5875	0.0376
bemkl_L_FP	0.4973 ±0.014	0.0469	0.5439	0.0863
bemkl_R_DIFP	0.5166 ±0.020	0.0137	0.4325	0.0266
bemkl_R_DI	0.5389 ±0.020	0.0183	0.5437	0.0354
bemkl_R_FP	0.4993 ±0.017	0.0101	0.3000	0.0195
RF_DIFP	0.4985 ±0.013	0.3791	0.4778	0.4228
RF_DI	0.5198 ±0.015	0.2159	0.5313	0.3070
RF_FP	0.4981 ±0.014	0.4402	0.4790	0.4588
NN_DIFP	0.5360 ±0.012	0.0163	0.5150	0.0315
NN_DI	0.5702 ±0.010	0.0192	0.6100	0.0373
NN_FP	0.4785 ±0.023	0.0104	0.3300	0.0202
LR_DIFP	0.4787 ±0.010	0.7588	0.4701	0.5805
LR_DI	0.4938 ±0.006	0.9229	0.4793	0.6309
LR_FP	0.4794 ±0.012	0.7534	0.4711	0.5797
SVM_J_DIFP	0.5041 ±0.014	0.4513	0.4864	0.4682
SVM_J_DI	0.5025 ±0.019	0.4952	0.4853	0.4902
SVM_J_FP	0.4976 ±0.019	0.4390	0.4784	0.4578
SVM_L_DIFP	0.5133 ±0.009	0.4521	0.4977	0.4738
SVM_L_DI	0.5333 ±0.016	0.3621	0.5464	0.4356
SVM_L_FP	0.4839 ±0.011	0.4796	0.4656	0.4725
SVM_R_DIFP	0.5067 ±0.015	0.2663	0.4870	0.3443
SVM_R_DI	0.5336 ±0.007	0.2187	0.5605	0.3146
SVM_R_FP	0.4719 ±0.014	0.3606	0.4411	0.3968
LDA_DIFP	0.5195 ±0.011	0.4564	0.5055	0.4797
LDA_DI	0.5342 ±0.012	0.4061	0.5331	0.4610
LDA_FP	0.5002 ±0.011	0.4896	0.4824	0.4860
NB_DIFP	0.5005 ±0.002	0.7976	0.4844	0.6028
NB_DI	0.5015 ±0.001	0.9229	0.4949	0.6443
NB_FP	0.4872 ±0.013	0.8009	0.4734	0.5950
RANDOM_P_DIFP	0.4971 ±0.028	0.4989	0.4793	0.4889
RANDOM_P_DI	0.4971 ±0.028	0.4989	0.4793	0.4889
RANDOM_P_FP	0.4971 ±0.028	0.4989	0.4793	0.4889

Continuation of Table 3				
Models	Accuracy	Precision	Recall	F-measure
RANDOM_W_DIFP	0.5029 ±0.028	0.5011	0.4856	0.4932
RANDOM_W_DI	0.5029 ±0.028	0.5011	0.4856	0.4932
RANDOM_W_FP	0.5029 ±0.028	0.5011	0.4856	0.4932
Weakness				
Models	Accuracy	Precision	Recall	F-measure
BEMKL_J_DIFP	0.5761 ±0.019	0.0148	0.4130	0.0286
BEMKL_J_DI	0.5529 ±0.017	0.0114	0.3130	0.0220
BEMKL_J_FP	0.4823 ±0.021	0.0211	0.4543	0.0404
BEMKL_L_DIFP	0.4952 ±0.018	0.0221	0.5957	0.0425
BEMKL_L_DI	0.5070 ±0.014	0.0106	0.2652	0.0204
BEMKL_L_FP	0.4948 ±0.024	0.0323	0.5609	0.0612
BEMKL_R_DIFP	0.5471 ±0.024	0.0200	0.5413	0.0386
BEMKL_R_DI	0.5518 ±0.015	0.0117	0.2851	0.0225
BEMKL_R_FP	0.4650 ±0.019	0.0177	0.3928	0.0338
RF_DIFP	0.5094 ±0.022	0.5637	0.5271	0.5448
RF_DI	0.5494 ±0.013	0.6423	0.5498	0.5924
RF_FP	0.5075 ±0.021	0.5453	0.5252	0.5351
NN_DIFP	0.4155 ±0.027	0.0108	0.3000	0.0208
NN_DI	0.4684 ±0.018	0.0144	0.3978	0.0277
NN_FP	0.4516 ±0.020	0.0164	0.4565	0.0317
LR_DIFP	0.5076 ±0.012	0.8253	0.5258	0.6423
LR_DI	0.5557 ±0.011	0.8928	0.5539	0.6837
LR_FP	0.4856 ±0.007	0.8482	0.5104	0.6373
SVM_J_DIFP	0.4913 ±0.015	0.5392	0.5116	0.5250
SVM_J_DI	0.5352 ±0.024	0.7666	0.5410	0.6344
SVM_J_FP	0.4841 ±0.015	0.5186	0.5025	0.5104
SVM_L_DIFP	0.4657 ±0.016	0.5002	0.4845	0.4922
SVM_L_DI	0.5225 ±0.006	0.5449	0.5302	0.5375
SVM_L_FP	0.4762 ±0.013	0.5125	0.4944	0.5033
SVM_R_DIFP	0.4502 ±0.016	0.6130	0.4812	0.5392
SVM_R_DI	0.4706 ±0.013	0.5463	0.5036	0.5241
SVM_R_FP	0.4730 ±0.018	0.5584	0.4981	0.5265
LDA_DIFP	0.4499 ±0.019	0.4859	0.4698	0.4777
LDA_DI	0.5088 ±0.011	0.5364	0.5205	0.5284
LDA_FP	0.4794 ±0.012	0.4913	0.4964	0.4938
NB_DIFP	0.5007 ±0.008	0.5442	0.5288	0.5364
NB_DI	0.5041 ±0.008	0.5810	0.5350	0.5570
NB_FP	0.4922 ±0.015	0.7181	0.5146	0.5995
RANDOM_P_DIFP	0.4769 ±0.048	0.4708	0.4912	0.4807
RANDOM_P_DI	0.4769 ±0.048	0.4708	0.4912	0.4807
RANDOM_P_FP	0.4769 ±0.048	0.4708	0.4912	0.4807
RANDOM_W_DIFP	0.5231 ±0.048	0.5292	0.5383	0.5338
RANDOM_W_DI	0.5231 ±0.048	0.5292	0.5383	0.5338
RANDOM_W_FP	0.5231 ±0.048	0.5292	0.5383	0.5338
Diarrhea				
Models	Accuracy	Precision	Recall	F-measure
BEMKL_J_DIFP	0.6529 ±0.014	0.0185	0.5900	0.0359
BEMKL_J_DI	0.6362 ±0.009	0.0167	0.5250	0.0323
BEMKL_J_FP	0.5650 ±0.025	0.0230	0.5740	0.0441
BEMKL_L_DIFP	0.5997 ±0.018	0.0272	0.8025	0.0526
BEMKL_L_DI	0.5940 ±0.011	0.0267	0.6863	0.0514
BEMKL_L_FP	0.5833 ±0.027	0.0423	0.7531	0.0800
BEMKL_R_DIFP	0.6251 ±0.019	0.0217	0.6850	0.0421
BEMKL_R_DI	0.6015 ±0.016	0.0168	0.5125	0.0326
BEMKL_R_FP	0.5709 ±0.023	0.0203	0.5076	0.0390
RF_DIFP	0.5705 ±0.014	0.5475	0.5800	0.5633
RF_DI	0.5602 ±0.006	0.2795	0.6604	0.3928
RF_FP	0.5609 ±0.013	0.5891	0.5635	0.5760
NN_DIFP	0.6266 ±0.014	0.0249	0.7900	0.0482
NN_DI	0.6054 ±0.012	0.0217	0.6900	0.0421
NN_FP	0.5558 ±0.025	0.0192	0.6100	0.0372
LR_DIFP	0.5180 ±0.024	0.7225	0.5260	0.6088
LR_DI	0.5272 ±0.010	0.8913	0.5237	0.6597
LR_FP	0.5209 ±0.021	0.7361	0.5278	0.6148
SVM_J_DIFP	0.5659 ±0.011	0.5991	0.5700	0.5842
SVM_J_DI	0.5041 ±0.014	0.5949	0.5102	0.5493
SVM_J_FP	0.5157 ±0.014	0.5983	0.5211	0.5570
SVM_L_DIFP	0.5687 ±0.012	0.5317	0.5807	0.5551
SVM_L_DI	0.5719 ±0.006	0.4322	0.6325	0.5135
SVM_L_FP	0.5479 ±0.021	0.5614	0.5524	0.5569
SVM_R_DIFP	0.6126 ±0.012	0.4809	0.6647	0.5581
SVM_R_DI	0.6050 ±0.006	0.3354	0.7320	0.4600
SVM_R_FP	0.5464 ±0.017	0.5920	0.5502	0.5703
LDA_DIFP	0.5682 ±0.019	0.5298	0.5802	0.5539
LDA_DI	0.5516 ±0.012	0.4656	0.5814	0.5171
LDA_FP	0.5427 ±0.013	0.5431	0.5499	0.5464
NB_DIFP	0.5002 ±0.000	0.9906	0.5093	0.6727
NB_DI	0.5000 ±0.000	1.0000	0.5063	0.6723
NB_FP	0.5508 ±0.015	0.6538	0.5519	0.5985
RANDOM_P_DIFP	0.4989 ±0.045	0.5165	0.5054	0.5109
RANDOM_P_DI	0.4989 ±0.045	0.5165	0.5054	0.5109
RANDOM_P_FP	0.4989 ±0.045	0.5165	0.5054	0.5109
RANDOM_W_DIFP	0.5011 ±0.045	0.4835	0.5081	0.4955
RANDOM_W_DI	0.5011 ±0.045	0.4835	0.5081	0.4955
RANDOM_W_FP	0.5011 ±0.045	0.4835	0.5081	0.4955
Rashes				
Models	Accuracy	Precision	Recall	F-measure
BEMKL_J_DIFP	0.5924 ±0.012	0.0183	0.5800	0.0354
BEMKL_J_DI	0.5500 ±0.013	0.0094	0.2950	0.0183
BEMKL_J_FP	0.5672 ±0.021	0.0179	0.4279	0.0344
BEMKL_L_DIFP	0.5447 ±0.013	0.0216	0.6292	0.0418

Continuation of Table 3

Models	Accuracy	Precision	Recall	F-measure
BEMKL_L_DI	0.5393 ±0.017	0.0192	0.5008	0.0370
BEMKL_L_FP	0.5141 ±0.024	0.0445	0.4504	0.0811
BEMKL_R_DIFP	0.5823 ±0.009	0.0216	0.6650	0.0418
BEMKL_R_DI	0.5455 ±0.009	0.0167	0.5030	0.0323
BEMKL_R_FP	0.5663 ±0.021	0.0142	0.3985	0.0274
RF_DIFP	0.5851 ±0.005	0.5630	0.5881	0.5753
RF_DI	0.5282 ±0.008	0.3345	0.5552	0.4175
RF_FP	0.5860 ±0.010	0.5607	0.5892	0.5746
NN_DIFP	0.5968 ±0.014	0.0186	0.5900	0.0360
NN_DI	0.5452 ±0.018	0.0175	0.5550	0.0339
NN_FP	0.5778 ±0.018	0.0131	0.4150	0.0254
LR_DIFP	0.5169 ±0.010	0.7876	0.5142	0.6222
LR_DI	0.4955 ±0.004	0.9366	0.4962	0.6487
LR_FP	0.5198 ±0.011	0.7803	0.5176	0.6224
SVM_J_DIFP	0.5525 ±0.020	0.5295	0.5544	0.5417
SVM_J_DI	0.5010 ±0.021	0.4459	0.5008	0.4717
SVM_J_FP	0.5251 ±0.011	0.4913	0.5238	0.5070
SVM_L_DIFP	0.5722 ±0.012	0.5497	0.5745	0.5618
SVM_L_DI	0.5217 ±0.008	0.3799	0.5379	0.4453
SVM_L_FP	0.5129 ±0.012	0.4947	0.5103	0.5024
SVM_R_DIFP	0.5427 ±0.013	0.5037	0.5429	0.5225
SVM_R_DI	0.5117 ±0.008	0.2627	0.5266	0.3505
SVM_R_FP	0.5807 ±0.011	0.5852	0.5793	0.5822
LDA_DIFP	0.5495 ±0.012	0.5289	0.5495	0.5390
LDA_DI	0.5248 ±0.013	0.4428	0.5351	0.4846
LDA_FP	0.5423 ±0.016	0.5379	0.5407	0.5393
NB_DIFP	0.5211 ±0.006	0.6035	0.5480	0.5744
NB_DI	0.5096 ±0.004	0.5796	0.5408	0.5595
NB_FP	0.5489 ±0.013	0.8035	0.5311	0.6395
RANDOM_P_DIFP	0.4883 ±0.054	0.4993	0.4852	0.4921
RANDOM_P_DI	0.4883 ±0.054	0.4993	0.4852	0.4921
RANDOM_P_FP	0.4883 ±0.054	0.4993	0.4852	0.4921
RANDOM_W_DIFP	0.5117 ±0.054	0.5007	0.5086	0.5046
RANDOM_W_DI	0.5117 ±0.054	0.5007	0.5086	0.5046
RANDOM_W_FP	0.5117 ±0.054	0.5007	0.5086	0.5046
Dermatitis				
Models	Accuracy	Precision	Recall	F-measure
BEMKL_J_DIFP	0.6100 ±0.010	0.0182	0.5800	0.0353
BEMKL_J_DI	0.5494 ±0.014	0.0096	0.3025	0.0186
BEMKL_J_FP	0.6013 ±0.016	0.0193	0.4342	0.0370
BEMKL_L_DIFP	0.5380 ±0.022	0.0177	0.5075	0.0343
BEMKL_L_DI	0.5306 ±0.014	0.0158	0.3971	0.0305
BEMKL_L_FP	0.5276 ±0.020	0.0403	0.4498	0.0740
BEMKL_R_DIFP	0.5979 ±0.008	0.0209	0.6275	0.0404
BEMKL_R_DI	0.5535 ±0.009	0.0190	0.4800	0.0366
BEMKL_R_FP	0.5917 ±0.016	0.0170	0.4108	0.0326
RF_DIFP	0.6032 ±0.013	0.6300	0.6129	0.6213
RF_DI	0.5384 ±0.011	0.3787	0.5858	0.4600
RF_FP	0.6075 ±0.010	0.6295	0.6182	0.6238
NN_DIFP	0.5829 ±0.019	0.0176	0.5600	0.0341
NN_DI	0.5128 ±0.009	0.0162	0.5150	0.0313
NN_FP	0.6038 ±0.014	0.0127	0.4050	0.0247
LR_DIFP	0.5205 ±0.018	0.8365	0.5316	0.6501
LR_DI	0.4932 ±0.007	0.9356	0.5118	0.6617
LR_FP	0.5215 ±0.015	0.8316	0.5328	0.6495
SVM_J_DIFP	0.4913 ±0.015	0.5392	0.5116	0.5250
SVM_J_DI	0.5352 ±0.024	0.7666	0.5410	0.6344
SVM_J_FP	0.4841 ±0.015	0.5186	0.5025	0.5104
SVM_L_DIFP	0.5587 ±0.016	0.5648	0.5717	0.5682
SVM_L_DI	0.5274 ±0.010	0.3989	0.5630	0.4669
SVM_L_FP	0.5252 ±0.016	0.5229	0.5385	0.5306
SVM_R_DIFP	0.5579 ±0.012	0.6380	0.5652	0.5994
SVM_R_DI	0.5392 ±0.012	0.3750	0.5972	0.4607
SVM_R_FP	0.5943 ±0.009	0.6719	0.5986	0.6332
LDA_DIFP	0.5449 ±0.016	0.5434	0.5590	0.5511
LDA_DI	0.5222 ±0.010	0.4468	0.5467	0.4918
LDA_FP	0.5464 ±0.010	0.5550	0.5594	0.5572
NB_DIFP	0.5149 ±0.007	0.5856	0.5576	0.5713
NB_DI	0.5019 ±0.004	0.6239	0.5430	0.5807
NB_FP	0.5558 ±0.013	0.7967	0.5539	0.6535
RANDOM_P_DIFP	0.5215 ±0.054	0.5449	0.5366	0.5407
RANDOM_P_DI	0.5215 ±0.054	0.5449	0.5366	0.5407
RANDOM_P_FP	0.5215 ±0.054	0.5449	0.5366	0.5407
RANDOM_W_DIFP	0.4785 ±0.054	0.4551	0.4933	0.4734
RANDOM_W_DI	0.4785 ±0.054	0.4551	0.4933	0.4734
RANDOM_W_FP	0.4785 ±0.054	0.4551	0.4933	0.4734
Vomiting				
Models	Accuracy	Precision	Recall	F-measure
BEMKL_J_DIFP	0.5847 ±0.015	0.0134	0.4421	0.0259
BEMKL_J_DI	0.5603 ±0.017	0.0095	0.3053	0.0185
BEMKL_J_FP	0.5675 ±0.020	0.0135	0.4053	0.0261
BEMKL_L_DIFP	0.5358 ±0.019	0.0195	0.5325	0.0377
BEMKL_L_DI	0.4881 ±0.017	0.0110	0.3184	0.0212
BEMKL_L_FP	0.5443 ±0.020	0.0518	0.5300	0.0944
BEMKL_R_DIFP	0.5680 ±0.021	0.0159	0.5263	0.0309
BEMKL_R_DI	0.5558 ±0.019	0.0086	0.2684	0.0166
BEMKL_R_FP	0.5573 ±0.020	0.0119	0.3675	0.0231
RF_DIFP	0.5625 ±0.011	0.3951	0.5567	0.4621
RF_DI	0.5157 ±0.006	0.1495	0.5109	0.2313
RF_FP	0.5636 ±0.010	0.4585	0.5457	0.4983
NN_DIFP	0.5059 ±0.015	0.0102	0.3368	0.0197

Continuation of Table 3

Models	Accuracy	Precision	Recall	F-measure
NN_DI	0.4466 ±0.012	0.0111	0.3684	0.0216
NN_FP	0.5726 ±0.020	0.0127	0.4211	0.0247
LR_DIFP	0.4757 ±0.017	0.6555	0.4534	0.5360
LR_DI	0.5078 ±0.008	0.9191	0.4710	0.6229
LR_FP	0.4712 ±0.018	0.6483	0.4475	0.5295
SVM_J_DIFP	0.5299 ±0.017	0.4751	0.4997	0.4871
SVM_J_DI	0.4811 ±0.013	0.3149	0.4349	0.3653
SVM_J_FP	0.5031 ±0.013	0.4247	0.4679	0.4452
SVM_L_DIFP	0.5162 ±0.009	0.4532	0.4840	0.4681
SVM_L_DI	0.5147 ±0.007	0.2708	0.4960	0.3503
SVM_L_FP	0.5413 ±0.013	0.4927	0.5112	0.5018
SVM_R_DIFP	0.5332 ±0.011	0.2528	0.5304	0.3424
SVM_R_DI	0.5349 ±0.007	0.1999	0.5450	0.2925
SVM_R_FP	0.5295 ±0.012	0.3492	0.5066	0.4134
LDA_DIFP	0.5191 ±0.015	0.4541	0.4875	0.4702
LDA_DI	0.4933 ±0.014	0.3221	0.4597	0.3788
LDA_FP	0.5290 ±0.012	0.4985	0.4962	0.4973
NB_DIFP	0.5119 ±0.005	0.7143	0.5510	0.6221
NB_DI	0.5082 ±0.004	0.8110	0.5262	0.6382
NB_FP	0.5328 ±0.012	0.6949	0.4949	0.5781
RANDOM_P_DIFP	0.4704 ±0.071	0.4858	0.4382	0.4608
RANDOM_P_DI	0.4704 ±0.071	0.4858	0.4382	0.4608
RANDOM_P_FP	0.4704 ±0.071	0.4858	0.4382	0.4608
RANDOM_W_DIFP	0.5296 ±0.071	0.5142	0.4986	0.5063
RANDOM_W_DI	0.5296 ±0.071	0.5142	0.4986	0.5063
RANDOM_W_FP	0.5296 ±0.071	0.5142	0.4986	0.5063

End of Table

Table 4: Statistical T-test for Abdominal Pain with different datasets

Drug Indications and Fingerprints **Abdominal Pain**

Models	BEMKL_J	BEMKL_R	BEMKL_L	RF	NN	LR	SVM_J	SVM_R	SVM_L	LDA	NB	R_P	R_W
BEMKL_J		2.98e-06	2.51e-10	5.81e-08	3.75e-08	1.11e-08	9.32e-09	5.34e-08	1.28e-08	1.13e-06	4.00e-07	8.54e-03	1.07e-02
BEMKL_R			2.58e-08	7.76e-04	3.90e-02	2.61e-04	1.68e-04	1.80e-02	4.07e-05	1.15e-02	4.97e-03	3.17e-01	4.29e-01
BEMKL_L				0.999	0.999	0.978	0.999	0.999	0.999	0.998	0.997	0.769	0.839
RF					1	0.0428	0.4894	1	1	1	1	1	1
NN							0.000258	0.006362	0.537715	0.052817	0.019989	0.771411	0.663947
LR								1	1	1	1	1	1
SVM_J								1	1	1	1	1	1
SVM_R									0.02625	0.07906	0.00217	0.69627	0.82687
SVM_L										1	0.984	1	1
LDA											0.447	0.893	0.981
NB												0.706	0.861
R_P													0.5424
R_W													

Drug Indications **Abdominal Pain**

Models	BEMKL_J	BEMKL_R	BEMKL_L	RF	NN	LR	SVM_J	SVM_R	SVM_L	LDA	NB	R_P	R_W
BEMKL_J		1.31e-05	6.58e-05	7.65e-07	1.60e-04	3.66e-08	1.38e-08	4.53e-08	2.90e-06	7.58e-07	3.85e-08	8.47e-03	8.68e-03
BEMKL_R			1.00e+00	4.47e-03	1.00e+00	6.30e-07	2.12e-06	1.53e-04	1.62e-02	1.42e-01	2.69e-06	1.46e-01	1.73e-01
BEMKL_L				1.41e-06	1.00e+00	1.33e-07	1.37e-09	1.69e-07	1.69e-06	8.64e-06	5.95e-09	4.11e-02	5.62e-02
RF					1.00e+00	4.48e-03	6.44e-04	1.00e+00	1.00e+00	1.00e+00	7.53e-05	7.78e-01	8.68e-01
NN						8.54e-07	4.53e-07	1.82e-06	1.06e-05	3.87e-05	9.33e-08	3.25e-02	4.53e-02
LR							1	1	1	1	0.773	1	1
SVM_J								1	1	1	0.00624	1	1
SVM_R									1.00e+00	1.00e+00	2.21e-05	4.00e-01	5.21e-01
SVM_L										1.00e+00	1.27e-07	1.90e-01	3.36e-01
LDA											6.55e-07	8.17e-02	1.51e-01
NB												0.707	0.864
R_P													0.5424
R_W													

Fingerprints **Abdominal Pain**

Models	BEMKL_J	BEMKL_R	BEMKL_L	RF	NN	LR	SVM_J	SVM_R	SVM_L	LDA	NB	R_P	R_W
BEMKL_J		1	1	1	1	1	1	1	1	1	1	1	1
BEMKL_R			1	1	1	1	1	1	1	1	1	1	1
BEMKL_L				1	1	1	1	1	1	1	1	1	1
RF					1	1	1	1	0.005341522	1	1	1	1
NN						5.94e-05	1.87e-05	9.85e-05	2.72e-06	4.59e-05	7.15e-05	4.98e-01	6.54e-01
LR							1	1	0.0275	1	1	1	1
SVM_J								1	0.00939	1	1	1	1
SVM_R									0.00195	1	1	1	1
SVM_L										1	1	1	1
LDA											1	1	1
NB												0.758	0.886
R_P													0.5424
R_W													

Table 5: Statistical T-test for Chronic Fatigue with different datasets

Drug Indications and Fingerprints **Chronic Fatigue**

Models	BEMKL_J	BEMKL_R	BEMKL_L	RF	NN	LR	SVM_J	SVM_R	SVM_L	LDA	NB	R_P	R_W
BEMKL_J		1.80e-05	3.62e-07	2.36e-05	4.50e-09	1.00e-05	3.86e-02	3.62e-07	8.26e-07	1.33e-07	2.43e-06	1.85e-02	2.16e-02
BEMKL_R			4.22e-04	4.55e-03	5.12e-07	1.97e-03	1.00e+00	1.28e-05	9.33e-05	4.03e-07	3.78e-04	2.16e-01	4.32e-01
BEMKL_L				1.00e+00	1.44e-05	1.00e+00	1.00e+00	1.04e-04	1.00e+00	1.13e-04	3.22e-01	1.00e+00	1.00e+00
RF					7.89e-04	1.00e+00	1.00e+00	3.86e-05	1.00e+00	7.38e-04	6.03e-01	1.00e+00	1.00e+00
NN						1	1	1	1	1	1	1	1
LR							1.000000	0.000261	1.000000	0.000445	0.596861	1.000000	1.000000
SVM_J								1.41e-05	5.79e-03	1.42e-05	2.65e-05	1.18e-01	1.27e-01
SVM_R									1	1	1	1	1
SVM_L										2.28e-05	1.81e-01	9.28e-01	1.00e+00
LDA											1	1	1
NB												0.779	1
R_P													0.5772
R_W													

Drug Indications **Chronic Fatigue**

Models	BEMKL_J	BEMKL_R	BEMKL_L	RF	NN	LR	SVM_J	SVM_R	SVM_L	LDA	NB	R_P	R_W
BEMKL_J		1.00e+00	1.06e-04	8.42e-04	1.53e-02	9.06e-06	4.18e-04	1.34e-05	1.97e-02	2.28e-03	3.30e-06	6.08e-02	7.73e-02
BEMKL_R			1.40e-03	4.93e-03	1.03e-01	7.74e-05	7.80e-04	1.45e-05	5.83e-01	4.54e-02	7.61e-06	4.91e-02	1.05e-01
BEMKL_L				1	1	0.02970	0.03124	0.00156	1	1	0.00941	0.70462	1
RF					1	6.12e-05	5.47e-03	2.69e-05	1	1	1.04e-04	3.33e-01	5.01e-01
NN						1.15e-05	3.60e-04	6.94e-05	1	1	2.03e-05	1.89e-01	2.03e-01
LR							0.20388	0.00226	1	1	0.04043	1	1
SVM_J								0.229	1	1	1	1	1
SVM_R									1	1	1	1	1
SVM_L										1.04e-03	6.95e-07	4.44e-02	5.48e-02
LDA											2.76e-06	7.16e-02	1.18e-01
NB												0.644	0.926
R_P													0.5772
R_W													

Fingerprints **Chronic Fatigue**

Models	BEMKL_J	BEMKL_R	BEMKL_L	RF	NN	LR	SVM_J	SVM_R	SVM_L	LDA	NB	R_P	R_W
BEMKL_J		0.154040	1	1	0.000235	1	1	1	1	1	0.636343	1	1
BEMKL_R			1	1	0.000766	1	1	1	1	1	1	1	1
BEMKL_L				1	0.000461	0.566674	1	0.534336	0.380812	0.237764	0.092634	1	1
RF					3.62e-05	1.82e-01	1	5.97e-01	9.93e-01	1.10e-01	1.70e-02	1	1
NN						1	1	1	1	1	1	1	1
LR							1	1	1	1	1	1	1
SVM_J								0.010853	0.106422	0.006536	0.000407	0.679682	0.977891
SVM_R									1	0.5218	0.0957	1	1
SVM_L										0.391	0.206	0.951	1
LDA											0.521	1	1
NB												0.808	1
R_P													0.5772
R_W													

Table 6: Statistical T-test for Dizziness with different datasets

Drug Indications and Fingerprints **Dizziness**

Models	BEMKL_J	BEMKL_R	BEMKL_L	RF	NN	LR	SVM_J	SVM_R	SVM_L	LDA	NB	R_P	R_W
BEMKL_J		1.80e-05	3.62e-07	2.36e-05	4.50e-09	1.00e-05	3.86e-02	3.62e-07	8.26e-07	1.33e-07	2.43e-06	1.85e-02	2.16e-02
BEMKL_R			4.22e-04	4.55e-03	5.12e-07	1.97e-03	1.00e+00	1.28e-05	9.33e-05	4.03e-07	3.78e-04	2.16e-01	4.32e-01
BEMKL_L				1.00e+00	1.44e-05	1.00e+00	1.00e+00	1.04e-04	1.00e+00	1.13e-04	3.22e-01	1.00e+00	1.00e+00
RF					7.89e-04	1.00e+00	1.00e+00	3.86e-05	1.00e+00	7.38e-04	6.03e-01	1.00e+00	1.00e+00
NN						1	1.00e+00	1	1.00e+00	1	1	1	1.00e+00
LR							1.00e+00	0.000261	1.00e+00	0.000445	0.596861	1	1.00e+00
SVM_J								1.41e-05	5.79e-03	1.42e-05	2.65e-05	1.18e-01	1.27e-01
SVM_R									1.00e+00	1	1	1	1
SVM_L										2.28e-05	1.81e-01	9.28e-01	1
LDA											1	1	1
NB												0.779	1
R_P													0.5772
R_W													

Drug Indications **Dizziness**

Models	BEMKL_J	BEMKL_R	BEMKL_L	RF	NN	LR	SVM_J	SVM_R	SVM_L	LDA	NB	R_P	R_W
BEMKL_J		4.78e-06	1.41e-04	3.34e-05	3.74e-03	4.70e-06	6.54e-03	2.15e-06	2.82e-05	5.46e-04	2.51e-06	1.64e-02	1.18e-02
BEMKL_R			1.000000	0.084364	1.000000	0.000547	1.000000	0.000137	0.268728	1.000000	0.000315	0.271524	0.461559
BEMKL_L				9.28e-03	1.000000	3.47e-05	1.000000	5.19e-05	5.15e-03	3.25e-01	2.29e-05	9.14e-02	9.21e-02
RF					1.000000	0.001641	1.000000	0.000198	1	1.000000	0.001190	0.375784	0.863986
NN						1.28e-05	2.64e-01	4.05e-06	4.09e-04	3.24e-02	1.26e-05	3.45e-02	7.68e-02
LR							1.000000	1	1		1	1	1
SVM_J								1.75e-05	9.48e-02	1.00e+00	1.73e-06	5.69e-02	1.74e-01
SVM_R									1	1	1	1	1
SVM_L										1	0.0012	0.1951	0.3730
LDA											4.93e-06	4.25e-02	8.87e-02
NB												0.804	0.969
R_P													0.5424
R_W													

Fingerprints **Dizziness**

Models	BEMKL_J	BEMKL_R	BEMKL_L	RF	NN	LR	SVM_J	SVM_R	SVM_L	LDA	NB	R_P	R_W
BEMKL_J		0.00144	0.00310	1	1	1	1	1	0.47789	1	1	1	1
BEMKL_R			0.0939	1	1	1	1	1	1	1	1	1	1
BEMKL_L				1	1	1	1	1	1	1	1	1	1
RF					1	1	1	1	0.000365	1	1	1	1
NN						1	1	1	0.0102	1	1	1	1
LR							1	1	2.06e-05	1	1	1	1
SVM_J								1	0.00641	1	1	1	1
SVM_R									3.78e-05	1	1	1	1
SVM_L										1	1	1	1
LDA											0.00331	1	1
NB												1	1
R_P													0.5424
R_W													

Table 7: Statistical T-test for Headache with different datasets

Drug Indications and Fingerprints **Headache**

Models	BEMKL_J	BEMKL_R	BEMKL_L	RF	NN	LR	SVM_J	SVM_R	SVM_L	LDA	NB	R_P	R_W
BEMKL_J		9.06e-08	6.12e-08	1.91e-09	2.69e-08	6.61e-09	4.41e-09	6.59e-11	1.43e-09	1.60e-10	1.97e-11	1.24e-05	2.31e-06
BEMKL_R			1.54e-06	3.24e-06	5.97e-05	4.79e-07	3.73e-08	1.11e-07	6.22e-08	6.04e-09	3.52e-10	2.14e-04	1.97e-05
BEMKL_L				1.00e+00	1.00e+00	4.69e-03	2.77e-02	1.00e+00	1.00e+00	5.56e-04	3.14e-06	6.45e-02	8.23e-03
RF					1.00e+00	1.14e-06	5.30e-05	1.45e-04	3.47e-03	1.18e-06	1.24e-09	6.86e-03	9.71e-04
NN						2.80e-05	1.74e-04	4.98e-05	1.68e-04	1.28e-05	1.12e-07	5.43e-03	5.75e-04
LR							1	1	1	1	0.0546	1	0.1822
SVM_J								1.00e+00	1.00e+00	2.82e-03	1.20e-07	3.72e-01	6.91e-03
SVM_R									1.00e+00	1.35e-04	6.49e-06	1.40e-01	5.91e-03
SVM_L										1.65e-04	9.49e-07	5.09e-02	2.57e-03
LDA											0.0327	1	0.0839
NB												1	0.203
R_P													0.1219
R_W													

Drug Indications **Headache**

Models	BEMKL_J	BEMKL_R	BEMKL_L	RF	NN	LR	SVM_J	SVM_R	SVM_L	LDA	NB	R_P	R_W
BEMKL_J		4.09e-08	2.86e-07	3.19e-08	1.58e-07	8.43e-08	2.25e-09	2.79e-10	5.31e-11	1.78e-11	2.94e-12	1.52e-05	4.33e-06
BEMKL_R			1.00e+00	8.02e-04	1.00e+00	4.95e-02	2.19e-07	3.80e-07	4.06e-11	1.52e-07	1.75e-10	5.82e-04	5.97e-05
BEMKL_L				2.80e-05	7.91e-06	2.15e-05	3.62e-08	2.90e-08	1.66e-09	6.71e-11	6.43e-11	6.54e-05	1.50e-05
RF					1	1	1.48e-06	3.44e-05	1.05e-08	1.17e-05	1.11e-09	7.60e-04	1.54e-04
NN						3.98e-01	4.45e-08	3.18e-07	4.34e-07	2.37e-06	8.69e-11	6.51e-04	4.57e-05
LR							5.21e-05	1.46e-03	1.36e-05	5.21e-05	2.41e-08	5.83e-04	2.72e-04
SVM_J								1	1	1	2.29e-05	7.74e-01	1.22e-02
SVM_R									2.53e-04	3.60e-01	1.95e-10	1.12e-02	4.34e-04
SVM_L										1	5.08e-08	4.35e-02	2.40e-03
LDA											2.13e-08	9.44e-03	6.67e-04
NB												1	0.222
R_P													0.1219
R_W													

Fingerprints **Headache**

Models	BEMKL_J	BEMKL_R	BEMKL_L	RF	NN	LR	SVM_J	SVM_R	SVM_L	LDA	NB	R_P	R_W
BEMKL_J		7.09e-02	7.00e-03	1.00e+00	6.85e-02	5.77e-05	1.39e-05	1.00e+00	1.87e-06	9.99e-04	6.82e-08	2.09e-01	5.30e-03
BEMKL_R			8.05e-02	1.00e+00	9.62e-01	6.45e-04	4.13e-05	1.00e+00	6.01e-05	4.29e-03	2.63e-06	4.04e-01	9.40e-03
BEMKL_L				1.00e+00	1	0.1140	0.1687	1.00e+00	0.0109	1	0.0429	1	0.1575
RF					1.12e-02	1.68e-05	3.74e-06	2.08e-02	9.61e-07	1.28e-05	1.22e-06	2.97e-02	2.08e-03
NN						0.01167	0.00637	1.00e+00	0.00017	0.25371	0.00117	0.67150	0.01971
LR							1	1.00e+00	1	1	1	1	0.35
SVM_J								1	1	1	1	1	0.173
SVM_R									2.28e-06	1.50e-04	1.85e-06	8.10e-02	3.30e-03
SVM_L										1	1	1	0.237
LDA											0.00203	0.80533	0.03166
NB												1	0.102
R_P													0.1219
R_W													

Table 8: Statistical T-test for Nausea with different datasets

Drug Indications and Fingerprints **Nausea**

Models	BEMKL_J	BEMKL_R	BEMKL_L	RF	NN	LR	SVM_J	SVM_R	SVM_L	LDA	NB	R_P	R_W
BEMKL_J		2.83e-07	1.18e-08	4.49e-06	1.21e-05	6.36e-07	1.74e-06	4.70e-07	2.13e-07	5.04e-05	1.63e-06	1.08e-04	1.22e-03
BEMKL_R			0.05501	0.06807	1.00000	0.00297	0.20147	0.34359	1.00000	1.00000	0.15448	0.13599	1
BEMKL_L				1.00000	1.00000	0.00763	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000
RF					1.00000	0.0371	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000
NN						1.41e-05	6.09e-04	4.15e-06	1.34e-04	6.28e-02	2.85e-05	4.60e-03	7.01e-02
LR							1	1.00000	1	1.00000	1	1	1.00000
SVM_J								1.00000	1	1.00000	1	1	1.00000
SVM_R									1	1.00000	0.521	0.725	1.00000
SVM_L										1.00000	0.00338	0.23360	0.64435
LDA											0.00153	0.05369	0.22284
NB												0.706	1.00000
R_P													0.6218
R_W													

Drug Indications **Nausea**

Models	BEMKL_J	BEMKL_R	BEMKL_L	RF	NN	LR	SVM_J	SVM_R	SVM_L	LDA	NB	R_P	R_W
BEMKL_J		4.98e-06	1.67e-08	4.03e-07	5.00e-02	8.47e-08	9.94e-06	1.72e-05	7.19e-07	3.17e-05	3.27e-07	3.58e-05	3.23e-04
BEMKL_R			1.00e+00	2.28e-02	1.00e+00	5.78e-05	1.11e-02	1.00e+00	1.00e+00	1.00e+00	1.19e-03	5.14e-03	1.23e-01
BEMKL_L				5.71e-04	1.00e+00	8.32e-06	1.45e-03	9.19e-01	2.75e-01	1.00e+00	7.61e-05	7.40e-03	3.81e-02
RF					1.00e+00	0.00118	0.04127	1	1.00e+00	1.00e+00	0.02012	0.34122	0.45235
NN						1.21e-08	1.21e-06	2.70e-05	5.18e-04	1.17e-04	6.46e-08	2.23e-04	1.77e-04
LR							1	1	1.00e+00	1	1	1	1
SVM_J									1.00e+00	1	1	1	1
SVM_R									1.00e+00	1	5.70e-07	1.46e-02	1.46e-02
SVM_L										1	0.000282	0.006498	0.059507
LDA											2.21e-05	7.41e-03	1.63e-02
NB												0.635	1
R_P													0.6218
R_W													

Fingerprints **Nausea**

Models	BEMKL_J	BEMKL_R	BEMKL_L	RF	NN	LR	SVM_J	SVM_R	SVM_L	LDA	NB	R_P	R_W
BEMKL_J		1	1	1	0.04301	0.17931	1	0.00865	0.15826	1	1	1	1
BEMKL_R			1	1	0.02290	0.02355	1	0.00505	0.05197	1	0.30041	1	1
BEMKL_L				1	0.1508	0.0863	1	0.0111	0.0221	1	0.6518	1	1
RF					0.013182	0.042796	1	0.000406	0.069130	1	0.162527	1	1
NN						1	1	0.694	1	1	1	1	1
LR							1	1	1	1	1	1	1
SVM_J								0.00291	0.08843	1	0.32494	1	1
SVM_R									1	1	1	1	1
SVM_L										1	1	1	1
LDA											0.121	1	1
NB												1	1
R_P													0.6218
R_W													

Table 9: Statistical T-test for Weakness with different datasets

Drug Indications and Fingerprints **Weakness**

Models	BEMKL_J	BEMKL_R	BEMKL_L	RF	NN	LR	SVM_J	SVM_R	SVM_L	LDA	NB	R_P	R_W
BEMKL_J		2.29e-03	4.08e-05	1.21e-04	1.31e-07	1.79e-04	6.04e-07	4.62e-09	3.70e-08	1.56e-07	2.18e-06	1.22e-03	5.92e-02
BEMKL_R			4.20e-04	2.35e-04	2.60e-08	2.14e-02	8.92e-06	2.46e-08	4.89e-07	1.63e-07	3.66e-04	9.11e-03	1.00e+00
BEMKL_L				1.00e+00	2.22e-06	1.00e+00	1.00e+00	7.75e-04	8.65e-04	4.59e-04	1.00e+00	1.00e+00	1.00e+00
RF					1.65e-07	1.00e+00	2.43e-02	4.47e-06	1.04e-04	1.94e-05	9.01e-01	1.59e-01	1.00e+00
NN						1	1.00e+00	1	1		1.00e+00	1	1.00e+00
LR							0.176499	0.000162	0.002021	0.000153	0.718182	0.325362	1.00e+00
SVM_J								2.65e-05	4.98e-04	2.05e-04	1.00e+00	1.00e+00	1.00e+00
SVM_R									1	1	1.00e+00	1	1.00e+00
SVM_L										0.0262	1.00e+00	1	1.00e+00
LDA											1.00e+00	1	1.00e+00
NB												0.18	1.00e+00
R_P													0.9182
R_W													

Drug Indications **Weakness**

Models	BEMKL_J	BEMKL_R	BEMKL_L	RF	NN	LR	SVM_J	SVM_R	SVM_L	LDA	NB	R_P	R_W
BEMKL_J		1.00e+00	1.32e-06	1.00e+00	6.98e-07	1.00e+00	3.59e-01	1.60e-05	9.18e-04	2.60e-04	1.64e-04	2.51e-03	7.94e-01
BEMKL_R			4.17e-06	1.00e+00	2.07e-06	1.00e+00	4.48e-01	6.48e-06	4.22e-04	6.99e-05	1.88e-05	5.07e-03	6.27e-01
BEMKL_L				1.00e+00	0.000596	1.00e+00	1	0.004722	1	1	1	0.355885	1
RF					1.54e-07	1.00e+00	9.90e-02	8.33e-08	1.49e-03	2.93e-04	1.87e-05	5.13e-03	6.18e-01
NN						1.00e+00	1		1	1	1	1	1
LR							1.83e-01	5.86e-07	1.95e-04	2.92e-05	1.31e-05	1.83e-03	2.63e-01
SVM_J								3.19e-05	5.35e-01	7.35e-02	1.04e-02	1.93e-02	1.00e+00
SVM_R									1	1	1	1	1.00e+00
SVM_L										0.000269	0.000204	0.030430	1.00e+00
LDA											0.338	0.104	1.00e+00
NB												0.136	1.00e+00
R_P													0.9182
R_W													

Fingerprints **Weakness**

Models	BEMKL_J	BEMKL_R	BEMKL_L	RF	NN	LR	SVM_J	SVM_R	SVM_L	LDA	NB	R_P	R_W
BEMKL_J		1.06e-06	1.00e+00	1.00e+00	1.20e-04	1.00e+00	1.00e+00	7.17e-01	9.75e-01	1.00e+00	1.00e+00	1.00e+00	1.00e+00
BEMKL_R			1.00e+00	1.00e+00	0.0304	1.00e+00	1.00e+00	1	1		1.00e+00	1.00e+00	1.00e+00
BEMKL_L				1.00e+00	0.00062	1.00e+00	1.00e+00	0.19279	0.05091	0.22849	1.00e+00	1.00e+00	1.00e+00
RF					4.19e-06	5.29e-02	2.75e-02	1.28e-05	1.12e-03	1.45e-03	5.13e-02	2.21e-01	1.00e+00
NN						1.00e+00	1.00e+00	1	1	1.00e+00	1.00e+00	1.00e+00	1.00e+00
LR							1.00e+00	0.238	0.543	0.752	1.00e+00	1.00e+00	1.00e+00
SVM_J								0.282	0.296	0.915	1.00e+00	1.00e+00	1.00e+00
SVM_R									1	1.00e+00	1.00e+00	1.00e+00	1.00e+00
SVM_L										1.00e+00	1.00e+00	1.00e+00	1.00e+00
LDA											1.00e+00	1.00e+00	1.00e+00
NB												0.253	1.00e+00
R_P													0.9182
R_W													

Table 10: Statistical T-test for Diarrhea with different datasets

Drug Indications and Fingerprints **Diarrhea**

Models	BEMKL_J	BEMKL_R	BEMKL_L	RF	NN	LR	SVM_J	SVM_R	SVM_L	LDA	NB	R_P	R_W
BEMKL_J		2.29e-03	4.08e-05	1.21e-04	1.31e-07	1.79e-04	6.04e-07	4.62e-09	3.70e-08	1.56e-07	2.18e-06	1.22e-03	5.92e-02
BEMKL_R			4.20e-04	2.35e-04	2.60e-08	2.14e-02	8.92e-06	2.46e-08	4.89e-07	1.63e-07	3.66e-04	9.11e-03	1.00e+00
BEMKL_L				1.00e+00	2.22e-06	1.00e+00	1.00e+00	7.75e-04	8.65e-04	4.59e-04	1.00e+00	1.00e+00	1.00e+00
RF					1.65e-07	1.00e+00	2.43e-02	4.47e-06	1.04e-04	1.94e-05	9.01e-01	1.59e-01	1.00e+00
NN						1	1.00e+00	1	1	1	1.00e+00	1	1.00e+00
LR							0.176499	0.000162	0.002021	0.000153	0.718182	0.325362	1.00e+00
SVM_J								2.65e-05	4.98e-04	2.05e-04	1.00e+00	1	1.00e+00
SVM_R									1	1	1.00e+00	1	1.00e+00
SVM_L										0.0262	1.00e+00	1	1.00e+00
LDA											1.00e+00	1	1.00e+00
NB												0.18	1.00e+00
R_P													0.9182
R_W													

Drug Indications **Diarrhea**

Models	BEMKL_J	BEMKL_R	BEMKL_L	RF	NN	LR	SVM_J	SVM_R	SVM_L	LDA	NB	R_P	R_W
BEMKL_J		1.68e-05	3.70e-07	4.58e-09	3.09e-05	2.84e-10	6.05e-09	2.38e-05	2.56e-07	7.78e-09	5.44e-11	6.40e-05	2.75e-05
BEMKL_R			7.38e-01	8.57e-05	1.00e+00	1.39e-09	8.26e-07	1.00e+00	4.52e-03	1.86e-05	5.03e-08	1.55e-03	1.05e-04
BEMKL_L				2.28e-05	1.00e+00	1.33e-07	3.99e-07	1.00e+00	3.25e-03	7.50e-06	5.54e-09	7.59e-04	6.78e-04
RF					1.00e+00	5.92e-05	5.61e-07	1.00e+00	1.00e+00	7.28e-02	2.34e-09	6.45e-03	1.98e-02
NN						2.88e-07	2.79e-07	1.00e+00	1.28e-04	3.80e-06	3.22e-09	1.80e-04	3.44e-04
LR							1.04e-02	1.00e+00	1.00e+00	1.00e+00	8.68e-05	4.64e-01	2.21e-01
SVM_J								1.00e+00	1	1	1	1	1
SVM_R									7.04e-06	3.52e-07	6.09e-12	1.37e-04	1.24e-04
SVM_L										1.61e-03	2.14e-10	1.31e-03	2.13e-03
LDA											3.98e-07	9.42e-03	1.29e-02
NB												0.942	1
R_P													0.5291
R_W													

Fingerprints **Diarrhea**

Models	BEMKL_J	BEMKL_R	BEMKL_L	RF	NN	LR	SVM_J	SVM_R	SVM_L	LDA	NB	R_P	R_W
BEMKL_J		1	1	1	1	0.00517	0.00016	0.06307	0.06590	0.04184	0.83410	0.05074	0.00416
BEMKL_R			1	0.180106	0.345805	0.002479	0.000054	0.006066	0.009759	0.008996	0.232544	0.025945	0.002063
BEMKL_L				2.59e-02	4.86e-02	5.10e-04	8.18e-05	6.70e-03	1.38e-04	2.17e-03	2.72e-02	4.10e-03	3.75e-03
RF					1	1.69e-03	4.06e-05	4.13e-02	7.94e-02	2.35e-03	4.86e-01	1.93e-02	8.56e-03
NN						0.00483	0.00819	0.38788	0.91460	0.30749	1	0.02698	0.03938
LR							1	1	1	1	1	0.762	0.799
SVM_J								1	1	1	1	1	0.775
SVM_R									1	1	1	0.0467	0.0305
SVM_L										0.8760	1	0.0215	0.0408
LDA											1	0.0326	0.0262
NB												0.00461	0.01486
R_P													0.5291
R_W													

Table 11: Statistical T-test for Rashes with different datasets

Drug Indications and Fingerprints **Rashes**

Models	BEMKL_J	BEMKL_R	BEMKL_L	RF	NN	LR	SVM_J	SVM_R	SVM_L	LDA	NB	R_P	R_W
BEMKL_J		2.13e-02	4.59e-05	4.74e-01	1.00e+00	3.69e-07	8.11e-04	3.78e-06	3.32e-02	2.07e-06	8.32e-07	2.10e-03	5.52e-03
BEMKL_R			1.48e-05	1.00e+00	1.00e+00	1.14e-06	7.76e-04	1.60e-05	3.22e-01	8.27e-05	2.81e-07	2.26e-03	1.77e-02
BEMKL_L				1.00e+00	1	0.00704	1	1	1	1	0.00311	0.07705	0.39668
RF					1.00e+00	1.32e-07	6.49e-04	1.12e-05	2.40e-02	5.81e-05	1.96e-09	1.68e-03	9.36e-03
NN						4.58e-07	1.93e-03	5.42e-06	1.01e-02	9.83e-07	7.17e-07	6.50e-04	4.63e-03
LR							1	1	1	1	1	0.544	1
SVM_J								0.36842	1	1	0.00369	0.02610	0.13725
SVM_R									1.00000	1.00000	0.00125	0.04335	0.24941
SVM_L										2.84e-03	1.88e-07	2.23e-03	1.47e-02
LDA											0.000446	0.012269	0.085120
NB												0.080	0.616
R_P													0.7423
R_W													

Drug Indications **Rashes**

Models	BEMKL_J	BEMKL_R	BEMKL_L	RF	NN	LR	SVM_J	SVM_R	SVM_L	LDA	NB	R_P	R_W
BEMKL_J		1.00e+00	2.16e-02	3.25e-03	1.00e+00	1.39e-06	2.94e-04	1.24e-04	5.09e-05	1.38e-04	2.87e-05	6.39e-02	2.54e-01
BEMKL_R			1.00e+00	5.59e-03	1.00e+00	1.10e-08	1.70e-03	2.93e-05	1.25e-04	2.27e-03	3.35e-06	5.27e-02	4.77e-01
BEMKL_L				0.35484	1.00000	0.00010	0.00196	0.00360	0.03713	0.03374	0.00197	0.12215	0.73438
RF					1	7.43e-07	4.95e-03	1.02e-02	5.35e-01	1	1.23e-03	1.66e-01	1
NN						2.39e-05	6.99e-04	1.12e-03	1.26e-02	1.83e-02	4.56e-04	4.02e-02	4.39e-01
LR							1	1	1	1	1	1	1
SVM_J								1	1	1	1	1	1
SVM_R									1	1	0.966	0.528	1
SVM_L										1	0.00516	0.23405	1
LDA											0.00882	0.10743	0.71778
NB												0.249	1
R_P													0.7423
R_W													

Fingerprints **Rashes**

Models	BEMKL_J	BEMKL_R	BEMKL_L	RF	NN	LR	SVM_J	SVM_R	SVM_L	LDA	NB	R_P	R_W
BEMKL_J		1.00e+00	2.68e-06	1.00e+00	1.00e+00	1.04e-03	7.35e-04	1.00e+00	1.96e-05	2.90e-03	3.13e-01	1.96e-02	6.21e-02
BEMKL_R			1.47e-05	1.00e+00	1.00e+00	9.94e-04	7.42e-04	1.00e+00	1.60e-05	4.81e-03	4.15e-01	1.65e-02	7.16e-02
BEMKL_L				1	1	1	1	1.00e+00	1	1	1	1	1
RF					4.63e-01	4.02e-08	2.94e-07	6.33e-01	2.70e-09	1.59e-07	2.47e-04	1.97e-03	8.33e-03
NN						4.76e-05	4.59e-05	1.00e+00	1.81e-07	9.23e-05	1.97e-02	5.59e-03	1.26e-02
LR							1	1.00e+00	0.336	1	1	0.427	1
SVM_J								1.00e+00	0.118	1	1	0.130	1
SVM_R									5.67e-08	2.88e-05	4.70e-04	1.60e-03	7.65e-03
SVM_L										1	1	0.476	1
LDA											1	0.0239	0.1812
NB												0.00719	0.06996
R_P													0.7423
R_W													

Table 12: Statistical T-test for Dermatitis with different datasets

Drug Indications and Fingerprints **Dermatitis**

Models	BEMKL_J	BEMKL_R	BEMKL_L	RF	NN	LR	SVM_J	SVM_R	SVM_L	LDA	NB	R_P	R_W
BEMKL_J		2.97e-02	3.00e-05	1.00e+00	1.06e-03	1.74e-06	2.61e-07	1.56e-06	4.11e-06	3.42e-06	1.78e-09	3.38e-03	2.90e-04
BEMKL_R			8.67e-05	1.00e+00	1.70e-01	1.84e-06	4.09e-08	1.77e-05	1.03e-04	2.08e-05	8.49e-10	1.14e-02	3.57e-04
BEMKL_L				1.0000	1.0000	0.2851	0.0012	1.0000	1.0000	1.0000	0.1145	1.0000	0.0290
RF					7.35e-02	4.99e-06	7.98e-08	1.05e-04	9.88e-06	1.77e-05	2.00e-07	8.46e-03	1.70e-04
NN						1.71e-04	2.91e-06	1.56e-02	8.37e-03	2.04e-03	7.61e-06	3.72e-02	1.02e-03
LR							0.0157	1.0000	1.0000	1.0000	1.0000	1.0000	0.2234
SVM_J								1	1	1	1	1	1
SVM_R									1.00e+00	3.96e-03	1.07e-06	2.26e-01	1.98e-03
SVM_L										0.20042	0.00011	0.09527	0.00552
LDA											0.000639	0.413453	0.003194
NB												1	0.0635
R_P													0.1219
R_W													

Drug Indications **Dermatitis**

Models	BEMKL_J	BEMKL_R	BEMKL_L	RF	NN	LR	SVM_J	SVM_R	SVM_L	LDA	NB	R_P	R_W
BEMKL_J		1.00e+00	2.68e-04	1.79e-01	9.80e-06	1.38e-05	1.00e+00	1.04e-01	7.57e-03	2.58e-04	4.58e-05	9.25e-01	1.94e-02
BEMKL_R			1.71e-03	9.56e-02	1.57e-05	4.90e-07	4.08e-01	2.12e-02	1.65e-03	5.37e-04	1.40e-07	5.43e-01	1.19e-02
BEMKL_L				1	0.005273	0.000173	1.00e+00	1	1	0.504003	0.001065	1	0.076984
RF					7.40e-04	1.63e-06	1.00e+00	1	2.50e-01	3.36e-02	2.83e-05	1	3.11e-02
NN						0.00562	1.00e+00	1	1	1	0.11156	1	0.41007
LR							1.00e+00	1	1	1	1	1	1
SVM_J								1	1	0.47776	0.00976	1	0.05298
SVM_R									0.062852	0.025155	0.000016	0.836921	0.021768
SVM_L										0.22682	0.00043	1	0.04997
LDA											0.00269	1	0.05800
NB												1	0.205
R_P													0.1219
R_W													

Fingerprints **Dermatitis**

Models	BEMKL_J	BEMKL_R	BEMKL_L	RF	NN	LR	SVM_J	SVM_R	SVM_L	LDA	NB	R_P	R_W
BEMKL_J		3.16e-02	1.79e-07	1	1	2.72e-06	5.91e-08	1	5.74e-07	4.15e-06	3.93e-04	1.41e-02	2.91e-04
BEMKL_R			7.16e-07	1	1	2.05e-05	2.00e-07	1	3.05e-07	9.45e-06	3.80e-03	2.86e-02	4.76e-04
BEMKL_L				1	1	1	0.00189	1	1	1	1	1.41e-02	0.09204
RF					1.00e+00	9.07e-07	5.82e-09	2.88e-02	4.73e-08	1.74e-07	3.19e-05	4.41e-03	1.65e-04
NN						4.33e-06	7.27e-09	6.98e-01	1.74e-06	5.18e-07	1.53e-05	1.50e-02	3.23e-05
LR							0.00239	1	1	1	1	1	0.21686
SVM_J								1	1	1	1	1	1
SVM_R									6.34e-07	5.44e-07	7.21e-05	3.98e-03	3.93e-04
SVM_L										1	1	1	0.0485
LDA											1	0.31123	0.00443
NB												0.124285	0.000598
R_P													0.1219
R_W													

Table 13: Statistical T-test for Vomiting with different datasets

Drug Indications and Fingerprints **Vomiting**													
Models	BEMKL_J	BEMKL_R	BEMKL_L	RF	NN	LR	SVM_J	SVM_R	SVM_L	LDA	NB	R_P	R_W
BEMKL_J		8.14e-03	5.13e-06	3.94e-03	3.86e-07	7.71e-08	3.54e-05	1.08e-05	5.41e-07	2.58e-07	5.20e-07	7.43e-03	1.90e-01
BEMKL_R			1.76e-03	1.00e+00	6.80e-05	1.72e-06	2.87e-03	2.04e-03	1.14e-04	1.21e-05	3.37e-05	1.25e-02	8.12e-01
BEMKL_L				1.00e+00	2.19e-03	3.65e-05	1.00e+00	1.00e+00	5.98e-03	7.04e-02	2.43e-02	1.41e-01	1.00e+00
RF					8.87e-07	4.90e-07	4.31e-04	7.11e-04	1.33e-06	2.57e-06	2.39e-06	2.03e-02	7.14e-01
NN						0.014	1.00e+00	1	1	1	1	0.808	1
LR							1.00e+00	1	1	1	1	1	1
SVM_J								1	0.00515	0.21590	0.05915	0.09373	1
SVM_R									0.000155	0.054224	0.002528	0.050047	1
SVM_L										1	0.649	0.173	1
LDA											0.254	0.107	1
NB												0.0955	1
R_P													0.8876
R_W													

Drug Indications **Vomiting**													
Models	BEMKL_J	BEMKL_R	BEMKL_L	RF	NN	LR	SVM_J	SVM_R	SVM_L	LDA	NB	R_P	R_W
BEMKL_J		1.00e+00	8.54e-11	6.78e-05	9.70e-08	1.14e-05	1.47e-05	1.20e-02	3.42e-05	3.17e-06	4.16e-05	3.25e-02	1.00e+00
BEMKL_R			3.53e-06	8.81e-04	4.14e-07	3.43e-04	2.00e-05	1.51e-02	7.30e-04	4.98e-05	9.14e-05	3.03e-02	1.00e+00
BEMKL_L				1	0.000324	1	1	1	1	1	1	1	1.00e+00
RF					8.24e-07	2.79e-02	1.50e-03	1.00e+00	1.00e+00	8.06e-03	1.45e-01	3.24e-01	1.00e+00
NN						1	1	1.00e+00	1.00e+00	1	1	1	1.00e+00
LR							0.00441	1.00e+00	1.00e+00	0.02655	1	0.51287	1.00e+00
SVM_J								1.00e+00	1.00e+00	1	1	1	1.00e+00
SVM_R									7.56e-04	7.11e-05	5.53e-06	4.64e-02	1.00e+00
SVM_L										9.38e-05	6.06e-02	2.01e-01	1.00e+00
LDA											1	0.56	1.00e+00
NB												0.129	1.00e+00
R_P													0.8876
R_W													

Fingerprints **Vomiting**													
Models	BEMKL_J	BEMKL_R	BEMKL_L	RF	NN	LR	SVM_J	SVM_R	SVM_L	LDA	NB	R_P	R_W
BEMKL_J		1.27e-02	1.61e-02	1.00e+00	1.00e+00	1.65e-06	8.24e-06	3.51e-04	1.38e-03	2.67e-03	1.56e-02	1.68e-02	8.34e-01
BEMKL_R			2.14e-01	1.00e+00	1.00e+00	7.90e-06	2.41e-05	6.17e-03	1.83e-02	1.58e-02	6.78e-02	2.93e-02	1
BEMKL_L				1.00e+00	1.00e+00	1.47e-05	7.64e-04	6.68e-02	1	2.17e-01	1	7.34e-02	1
RF					1.00e+00	4.70e-07	1.40e-08	1.16e-04	1.70e-03	3.32e-05	1.62e-03	1.74e-02	6.78e-01
NN						5.50e-06	5.69e-06	4.63e-06	1.15e-04	2.99e-04	7.52e-03	2.03e-02	1.85e-01
LR							1	1	1	1	1	1	1
SVM_J								1	1	1	1	0.608	1
SVM_R									1	1	1	0.0997	1
SVM_L										0.0788	0.4275	0.0356	1
LDA											1	0.0701	1
NB												0.0143	0.9013
R_P													0.8876
R_W													

[a] table ends

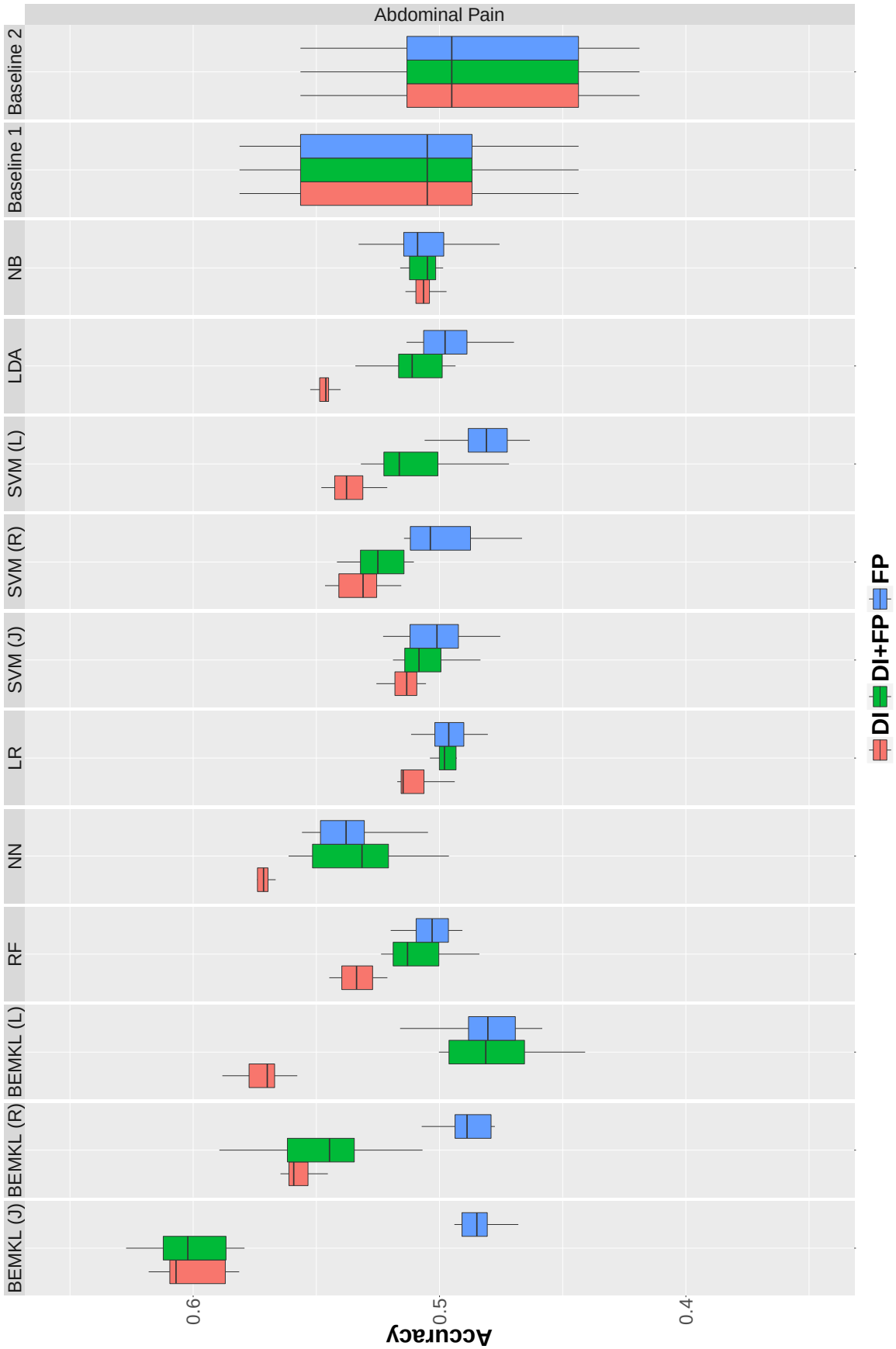

Figure 6: Performance measure for Side-effect : Abdominal Pain

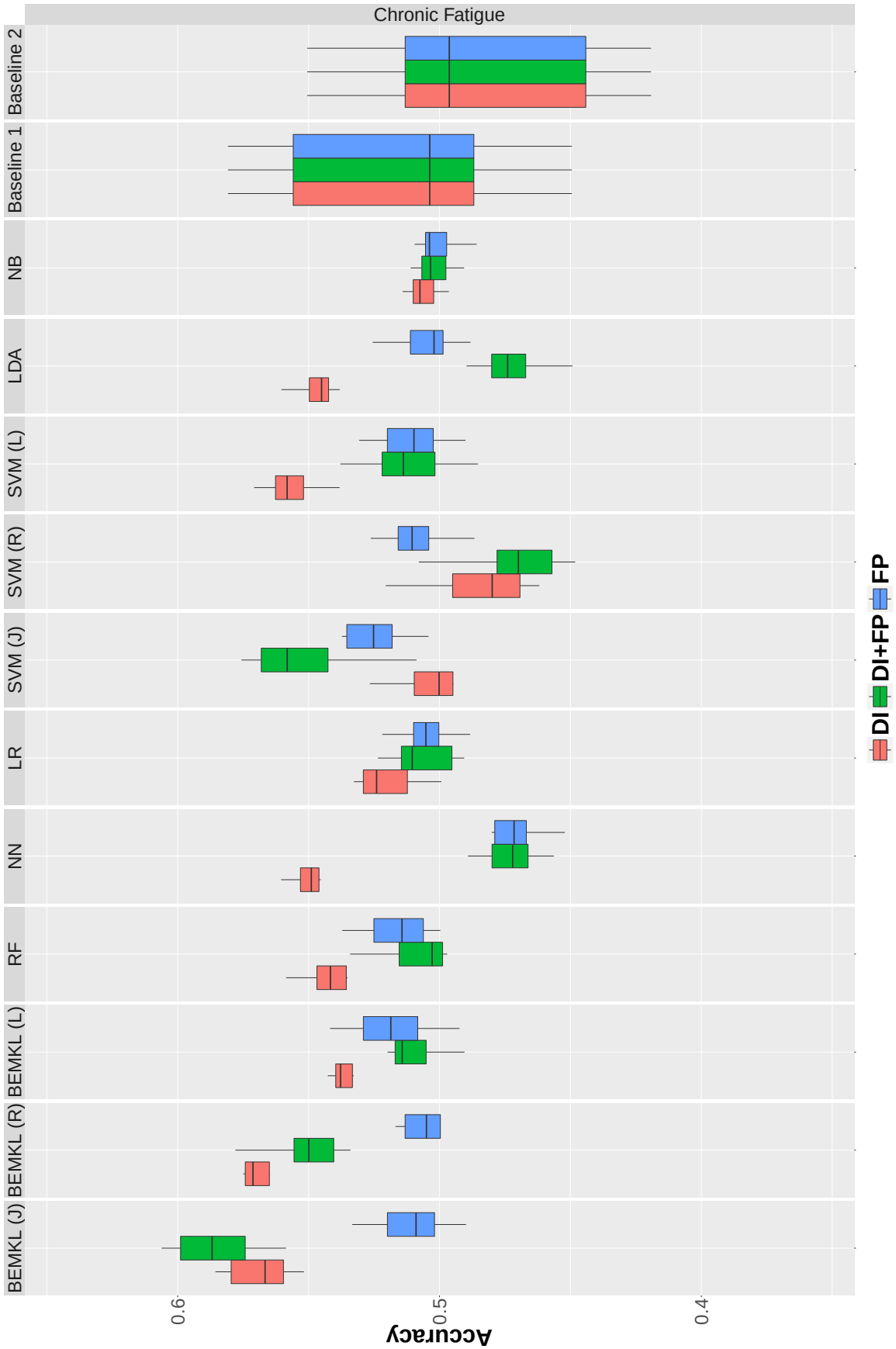

Figure 7: Performance measure for Side-effect : Chronic Fatigue

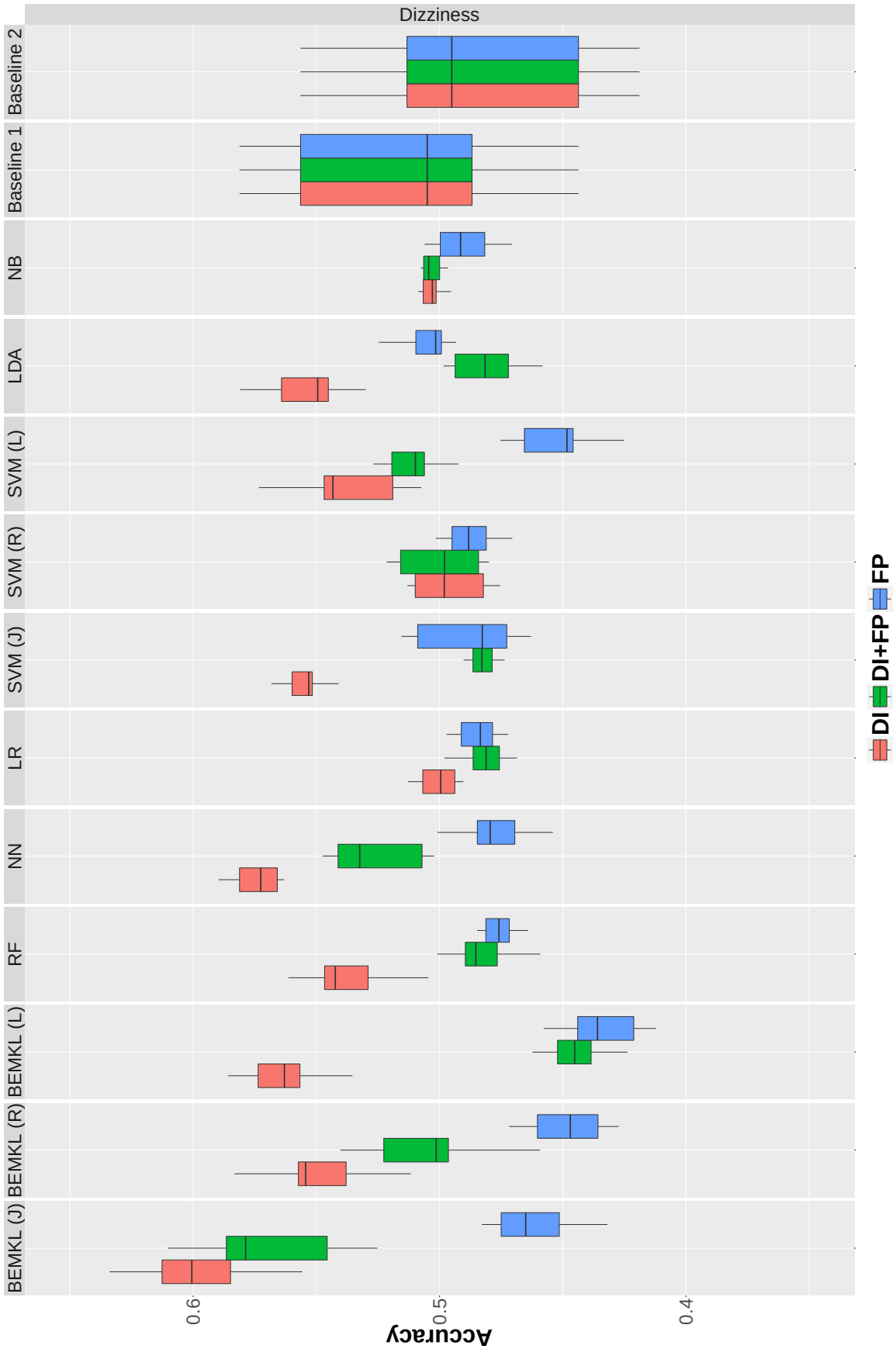

Figure 8: Performance measure for Side-effect : Dizziness

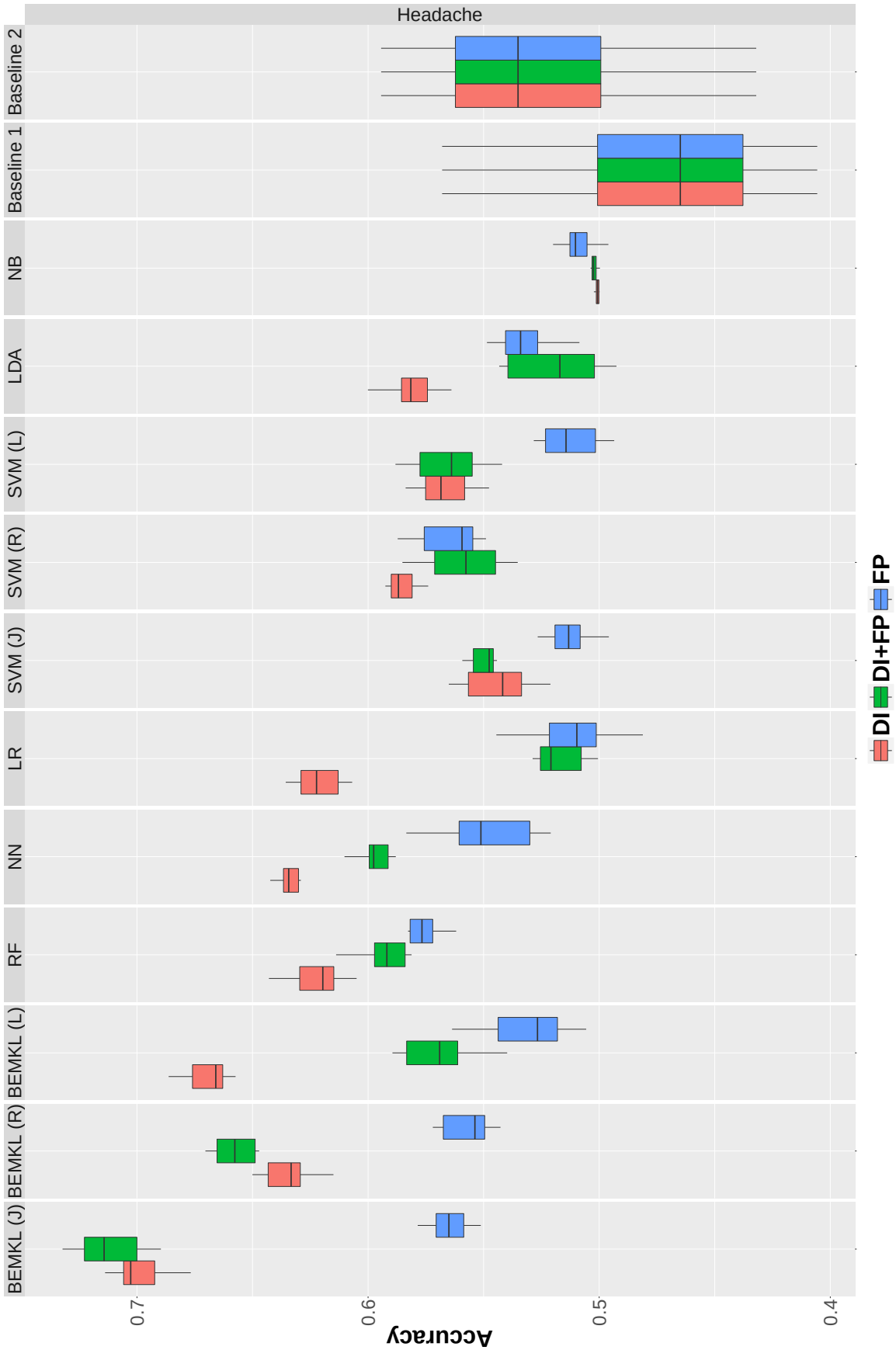

Figure 9: Performance measure for Side-effect : Headache

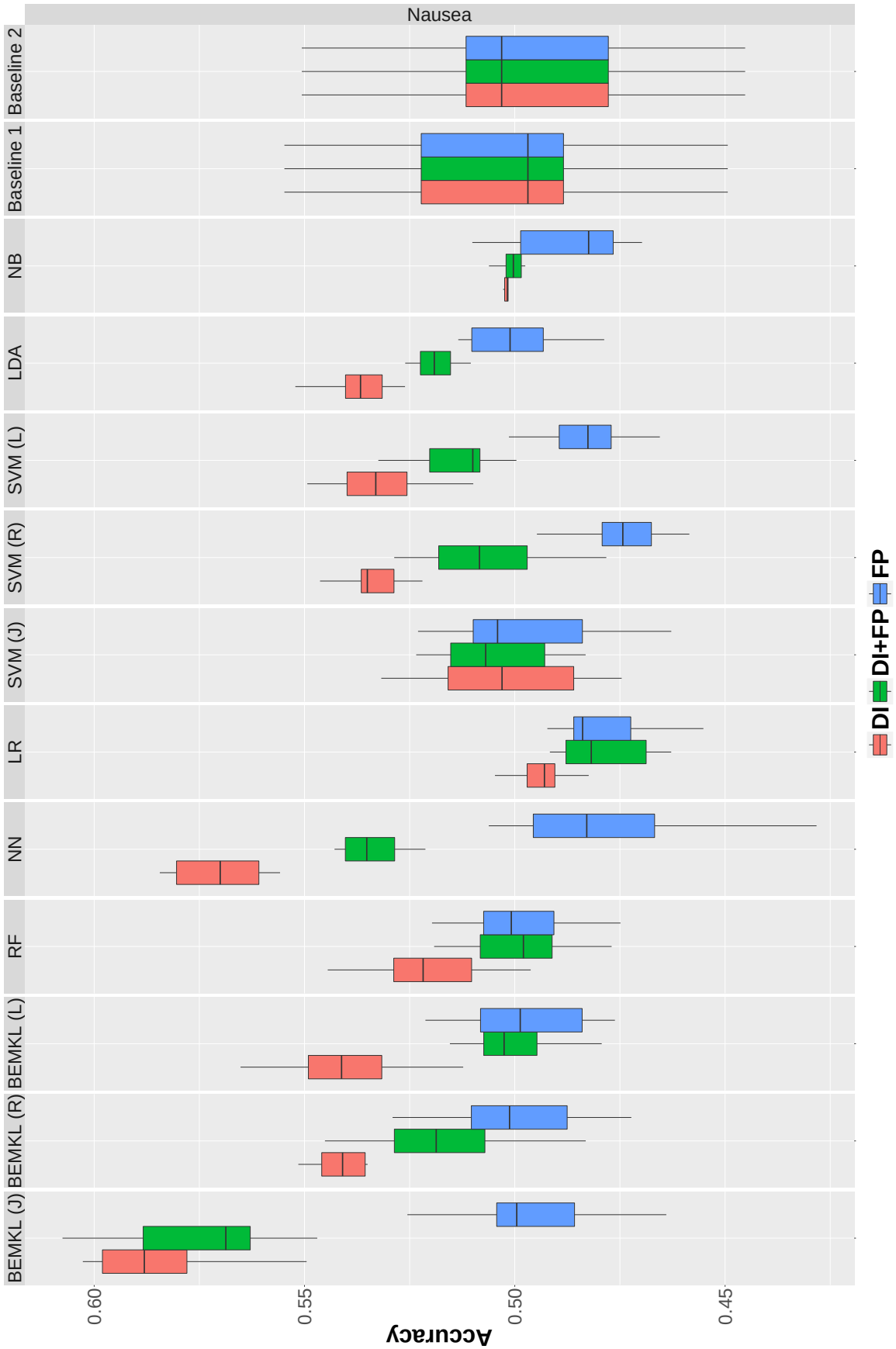

Figure 10: Performance measure for Side-effect : Nausea

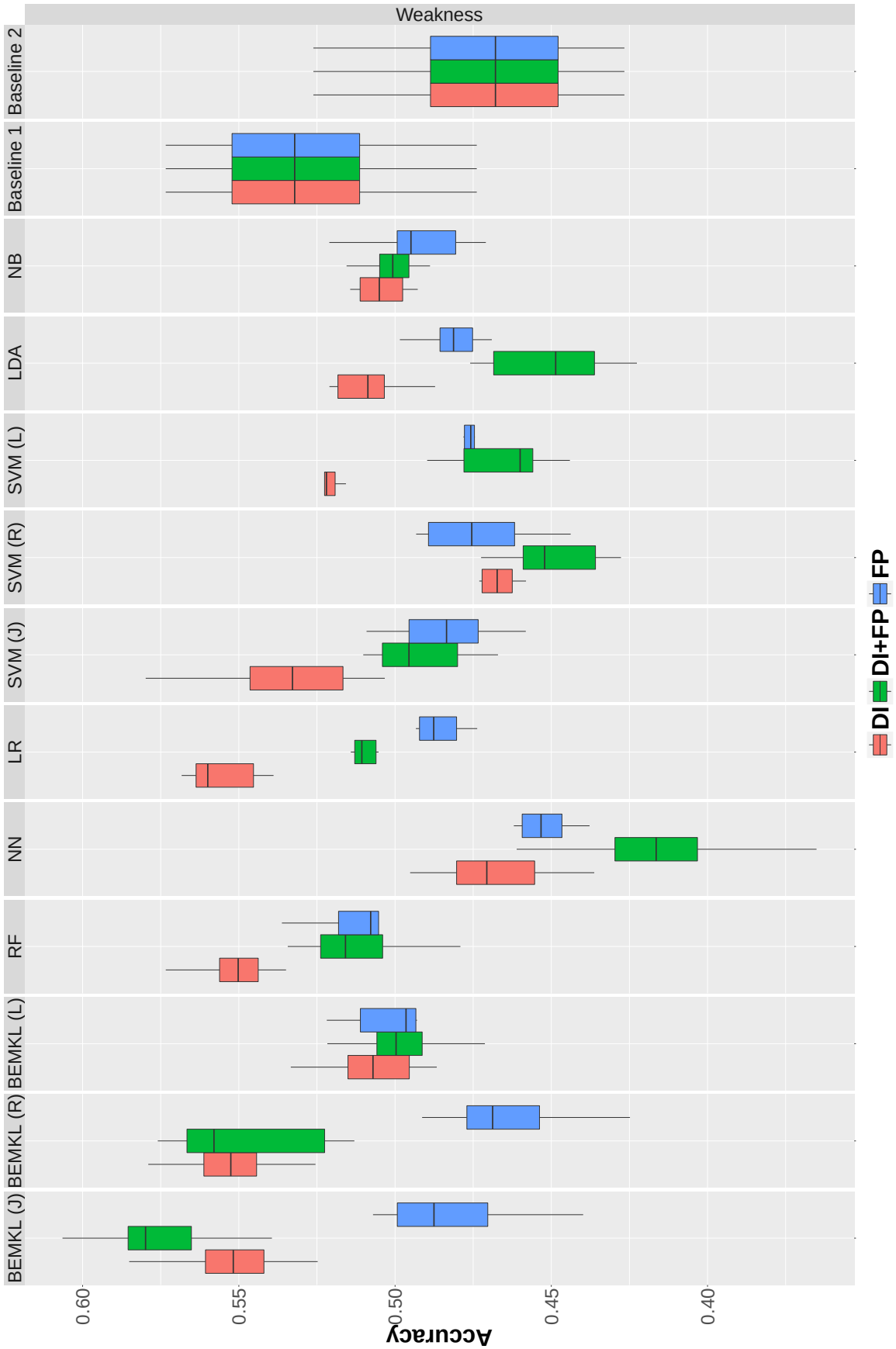

Figure 11: Performance measure for Side-effect : Weakness

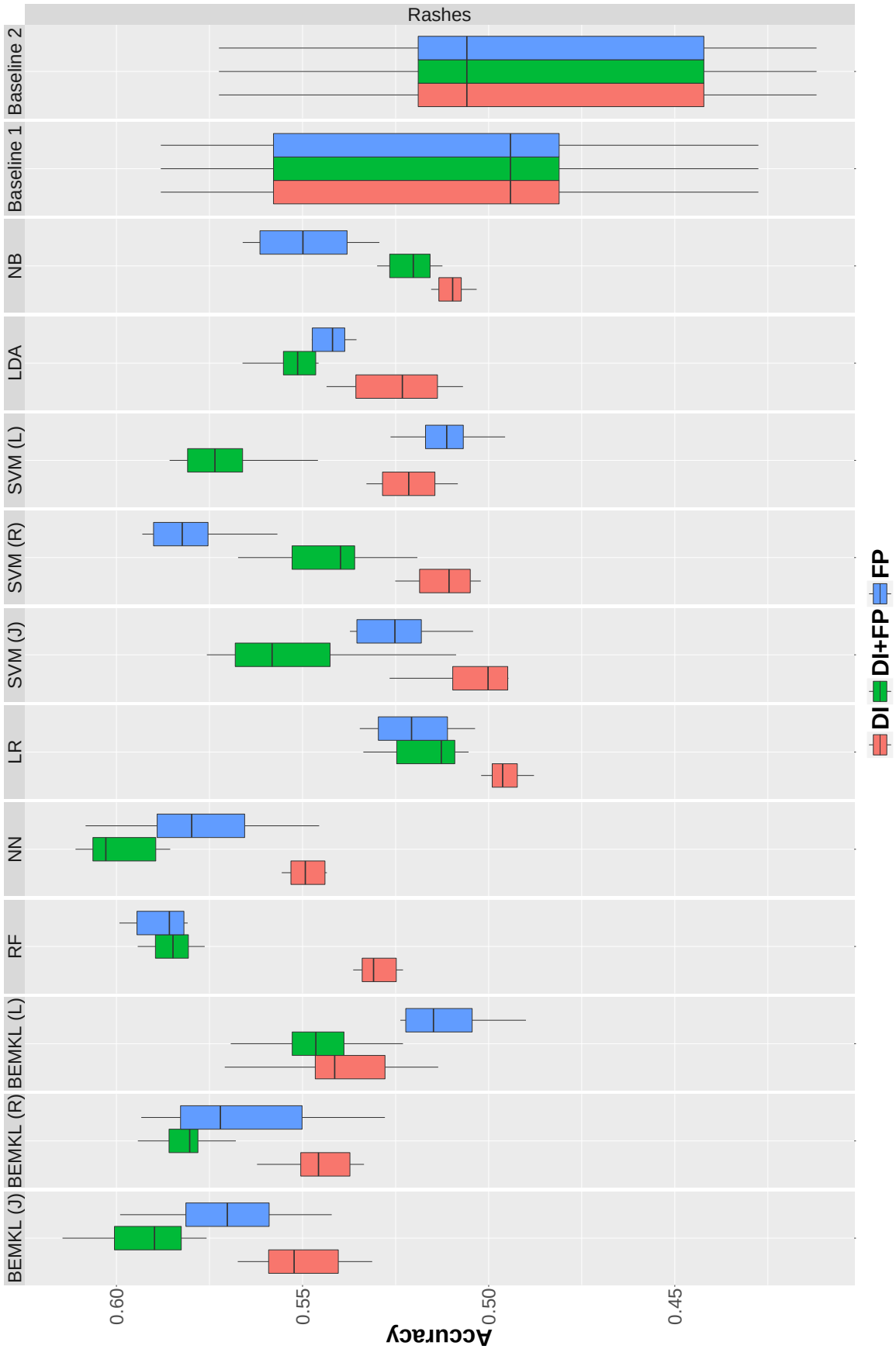

Figure 12: Performance measure for Side-effect : Rashes

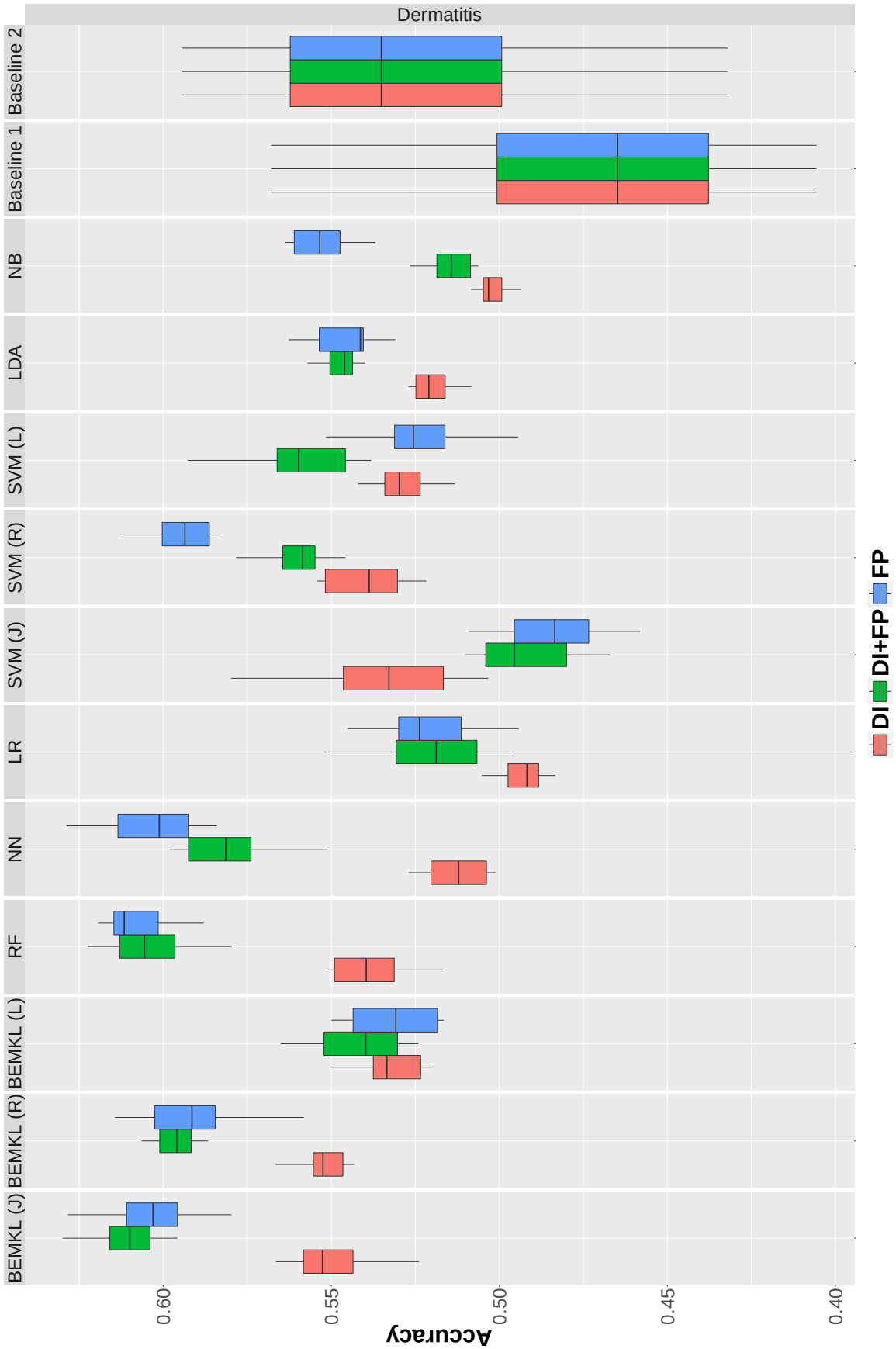

Figure 13: Performance measure for Side-effect : Dermatitis

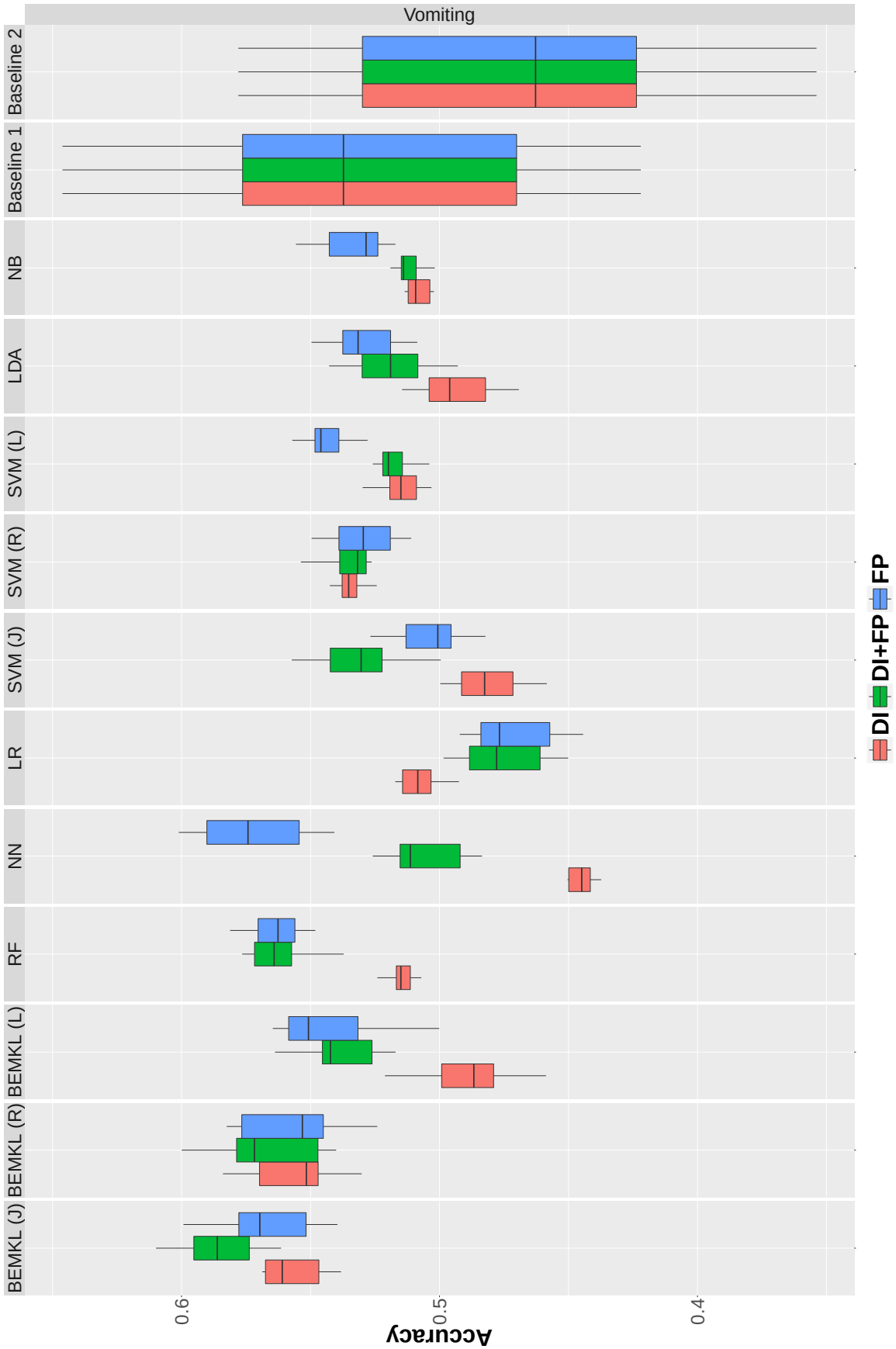

Figure 14: Performance measure for Side-effect : Vomiting